

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run On: July 3, 2003, 02:31:01 ; Search time 4275.04 seconds
(without alignments)
11062.367 Million cell updates/sec

Title: US-10-004-219B-2
Perfect score: 1625
Sequence: 1 gcttccagctgggtga.....ctgcaataaaatcagcagtc 1625

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :	GenEmbl.*
1: gb_ba.*	
2: gb_htg.*	
3: gb_in.*	
4: gb_om.*	
5: gb_ov.*	
6: gb_pat.*	
7: gb_ph.*	
8: gb_pl.*	
9: gb_pr.*	
10: gb_ro.*	
11: gb_sts.*	
12: gb_sy.*	
13: gb_un.*	
14: gb_vi.*	
15: em_ba.*	
16: em_fun.*	
17: em_hum.*	
18: em_in.*	
19: em_mu.*	
20: em_om.*	
21: em_or.*	
22: em_ov.*	
23: em_pat.*	
24: em_ph.*	
25: em_pl.*	
26: em_ro.*	
27: em_sts.*	
28: em_un.*	
29: em_vi.*	
30: em_htg_hum.*	
31: em_htg_inv.*	
32: em_htg_other.*	
33: em_htg_mus.*	
34: em_htg_pln.*	
35: em_htg_rod.*	
36: em_htg_mam.*	
37: em_htg_vrt.*	
38: em_sy.*	
39: em_htgo_hum.*	
40: em_htgo_mus.*	
41: em_htgo_other.*	

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1625	100.0	1625	9	AF290004	AF290004 Homo sapi
2	1266.2	77.9	1369	6	AX405989	AX405989 Sequence
3	1260.2	77.6	1354	9	AB025008	AB025008 Homo sapi
4	1095.2	67.4	1529	4	AB051629	AB051629 Bos tauru
5	1046	64.4	4250	9	AK098814	AK098814 Homo sapi
6	1045.6	64.3	1557	10	BC034548	BC034548 Mus muscu
7	1042.2	64.1	1538	10	BC011134	BC011134 Mus muscu
8	1041.8	64.1	1526	10	AF154571	AF154571 Mus muscu
9	1041.4	64.1	1530	10	AF290003	AF290003 Mus muscu
10	1037.2	63.8	1188	9	AB025009	AB025009 Homo sapi
11	701.8	43.2	1525	10	M94584	M94584 Mus musculu
12	697.2	42.9	1506	10	D87757	D87757 Mus musculu
13	689.4	42.4	1209	10	AY049765	AY049765 Mus muscu
14	687.8	42.3	1209	10	AY065557	AY065557 Mus muscu
15	483	29.7	1656	6	AR206042	AR206042 Sequence
16	483	29.7	1656	6	AR212139	AR212139 Sequence
17	483	29.7	1656	6	AX108752	AX108752 Sequence
18	481.4	29.6	1633	9	HS029615	U29615 Human chito
19	481.4	29.6	1643	6	AR172107	AR172107 Sequence
20	481	29.6	1636	6	AR206041	AR206041 Sequence
21	481	29.6	1636	6	AR212138	AR212138 Sequence
22	481	29.6	1636	6	AX108750	AX108750 Sequence
23	477.2	29.4	1713	6	AR172108	AR172108 Sequence
24	476.8	29.3	1710	9	HS062662	U62662 Homo sapien
25	457	28.1	994	10	MMU56900	U56900 Mus musculu
26	427	26.3	1562	5	BJA345054	AJ345054 Bufo japo
27	389.4	24.0	1599	9	AK055165	AK055165 Homo sapi
28	369	22.7	1433	6	AR042834	AR042834 Sequence
29	369	22.7	1434	9	HS058514	U58514 Human chiti
30	367.4	22.6	1418	9	HS049835	U49835 Human YKL-3
31	367.4	22.6	1449	9	BC011460	BC011460 Homo sapi
32	366	22.5	1500	9	HS058515	U58515 Human chiti
33	360	22.2	2237	9	MMU87259	U87259 Macaca mula
34	356.8	22.0	1406	9	AF484550	AF484550 Macaca ra
35	354.8	21.8	1526	6	AR042835	AR042835 Sequence
36	354.2	21.8	2198	9	HS009550	U09550 Human ovidu
37	354.2	21.8	2228	9	BABED0SG	M5903 Papio hamad
38	352.4	21.7	2525	10	MUS00GP	D32137 Mouse mrna
39	350.8	21.6	2504	6	E09046	E09046 cDNA encodi
40	350	21.5	2353	10	MAU15048	U15048 Mesocricetu
41	350	21.5	2387	10	HAM0GP	D32218 Hamster mrn
42	349.4	21.5	1149	6	E01501	E01501 cDNA encodi
43	349.4	21.5	1741	9	HUMH3G	M80927 Human glyco
44	348.4	21.4	91569	9	AL356387	AL356387 Human DNA
45	348.2	21.4	2034	4	OA016719	U16719 Ovis aries

ALIGNMENTS

RESULT 1	AF290004	AF290004	1625 bp	mrna	linear	PRI 27-FEB-2001
LOCUS	AF290004					
DEFINITION	Homo sapiens	acidic mammalian chitinase precursor, mRNA, complete cds.				
ACCESSION	AF290004					
VERSION	AF290004.1					
KEYWORDS						
SOURCE						
ORGANISM						
REFERENCE						
AUTHORS						

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1625)
Boot,R.G., Blommaert,E.F., Swart,E., Ghausharali-van der Vlugt,K.,
Bijl,N., Moe,C., Place,A., Aerts,J.M.


```
QY 1561 GTTCAGTCTCTTTTGTAGACATGTTGCCCTTACCTAAAGTCTCTCAATAAATCAG 1620
|||||
DB 1561 GTTCAGTCTCTTTTGTAGACATGTTGCCCTTACCTAAAGTCTCTCAATAAATCAG 1620

QY 1621 CAGTC 1625
|||||
DB 1621 CAGTC 1625

RESULT 2
LOCUS AX405989 1369 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 404 from Patent WO022660.
ACCESSION AX405989
VERSION AX405989.1 GI:21439417
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1. Tang, Y.T., Liu, C., Zhou, P., Asundi, V., Zhang, J., Zhao, Q.A., Ren, F.,
Xue, A.J., Yang, Y., Wehrman, T. and Drmanac, R.T.
Novel nucleic acids and polypeptides
Patent: WO 022660-A 404 21-MAR-2002;
HYSEQ, INC. (US)
FEATURES
source Location/Qualifiers
1..1369
/organism="Homo sapiens"
/db_xref="taxon:9606"
154..1260
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD35096.1"
/db_xref="GI:21439418"
/translation="MVSTPENRQTFITVIRKLFQYEFDFDWEYPSGRSPQDK
HLFTVLQVMEAFQEQAKQINKPRLMTAAVAAGISNIQSGYEIPQLSQYLDYIHYM
TYDLHSGEYTGENSEPLYKYPTDTSNAYLNVDYVMYKNDGAPAEKLIYGFYTYG
HNFILSNPTGIGAPTSGAGPAGYAKESGIWAYEICTFLKNGATGWDAPQVPI
AYQNVGVYDNIKSFIDIAKQWLKNKFGAMVWALDDDDFTGFCNKGKFLPILTLK
ALGLQASCTAPAPIETPITAPASGNGSGSSSGSGSGSGSGGFCVAVRANGLYPAN
NRNFAHCVNGVTYQONQAGLVFTTSCDCNWA"
BASE COUNT 327 a 393 c 341 g 308 t
ORIGIN

Query Match 77.9%; Score 1266.2; DB 6; Length 1369;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1274; Conservative. 0; Mismatches 13; Indels 0; Gaps 0;

QY 339 CTTTCAATGGCTGAAATAAGAACAGCCAGTGAAGTCTCTCGGCCATTGGAGGCT 398
|||||
DB 65 CTTTACTGAATTTGCAGCTCGGACAGCCAGCTGAAACTCTCTCGGCCATTGGAGGCT 124
|||||
QY 399 GGAACCTCGGAGTGGCCCTTTCTACTGCCATGTTTCTACTCCTTGAGAACCCGACAGCTT 458
|||||
DB 125 GGAACCTCGGAGTGGCCCTTTCTACTGCCATGTTTCTACTCCTTGAGAACCCGACAGCTT 184
|||||
QY 459 TCATCACCCTCAGTCATCAATTCCTCGGCCAGTATGAGTTTGACGGGCTGGACTTTGACT 518
|||||
DB 185 TCATCACCCTCAGTCATCAATTCCTCGGCCAGTATGAGTTTGACGGGCTGGACTTTGACT 244
|||||
QY 519 GSGAGTACCCTGGCTCTCGTGGAGCCCTCCTCAGGACAAGCATCTTTCACCTGTCTCTG 578
|||||
DB 245 GSGAGTACCCTGGCTCTCGTGGAGCCCTCCTCAGGACAAGCATCTTTCACCTGTCTCTG 304
|||||
QY 579 TGCAGGAATGGCTGAAGCTTTTGTAGCAGAGGCCAAGCAGATCAACAAGCCCGAGCTGA 638
|||||
DB 305 TGCAGGAATGGCTGAAGCTTTTGTAGCAGAGGCCAAGCAGATCAACAAGCCCGAGCTGA 364
|||||
QY 639 TGGTCACTGCTGAGTAGTCTGTCGATCTCCAAATATCAGTCTGCTATGAGATCCCC 698
|||||
DB 365 TGGTCACTGCTGAGTAGTCTGTCGATCTCCAAATATCAGTCTGCTATGAGATCCCC 424
|||||
```

```
QY 699 AACTGTACAGTACCTGGACTACATCCATGTGATGACCTACGACCTTCCATGGCTCTGGG 758
|||||
DB 475 AACTGTACAGTACCTGGACTACATCCATGTGATGACCTACGACCTTCCATGGCTCTGGG 484
|||||
QY 759 AGGGCTACAGTGGAGAGAACGCCCTCTTACAAATACCCGACTGACACCGGAGCAAGG 818
|||||
DB 485 AGGGCTACAGTGGAGAGAACGCCCTCTTACAAATACCCGACTGACACCGGAGCAAGG 544
|||||
QY 819 CCTACTCAATGTGATTTATGTCATGAACTACTGGAAGGACAAATGGAGCACCAGCTGAGA 878
|||||
DB 545 CCTACTCAATGTGATTTATGTCATGAACTACTGGAAGGACAAATGGAGCACCAGCTGAGA 604
|||||
QY 879 AGCTCATCGTTGGATTTCCTACTATGGACACAACCTTCTACTCTGAGCAACCCCTCCAACA 938
|||||
DB 605 AGCTCATCGTTGGATTTCCTACTATGGACACAACCTTCTACTCTGAGCAACCCCTCCAACA 664
|||||
QY 939 CTGGAATGGTGCCCTCCCACTCTGTGGTGGTCTGTGGGGCCCTATGCAAGGAGTCTG 998
|||||
DB 665 CTGGAATGGTGCCCTCCCACTCTGTGGTGGTCTGTGGGGCCCTATGCAAGGAGTCTG 724
|||||
QY 999 GGATCTGGGCTTACTACGAGATCTGTACCTTCTGAAAAATGGAGCACTCAGGGATGG 1058
|||||
DB 725 GGATCTGGGCTTACTACGAGATCTGTACCTTCTGAAAAATGGAGCACTCAGGGATGG 784
|||||
QY 1059 ATGCCCTCTCAGGAAGTGCCTTATGCTATCAGGGCAATGTGGTGTGCTATGACAACA 1118
|||||
DB 785 ATGCCCTCTCAGGAAGTGCCTTATGCTATCAGGGCAATGTGGTGTGCTATGACAACA 844
|||||
QY 1119 TCAAGAGCTTCGATTTAAGGCTCAATGGCTTAAGCACAAACAAATTTGGAGGGCCATGG 1178
|||||
DB 845 TCAAGAGCTTCGATTTAAGGCTCAATGGCTTAAGCACAAACAAATTTGGAGGGCCATGG 904
|||||
QY 1179 TCTGGGCCATTGATCTGGATGACTTCACCTGGGCACTTTCTGCAACAGGCAAGTTTCCCC 1238
|||||
DB 905 TCTGGGCCATTGATCTGGATGACTTCACCTGGGCACTTTCTGCAACAGGCAAGTTTCCCC 964
|||||
QY 1239 TAATCTCCACCCTGAAGAAGGCCCTCGGCTCGAGAGTGCAGAGTTCAGCGCTCCAGCTC 1298
|||||
DB 965 TAATCTCCACCCTGAAGAAGGCCCTCGGCTCGAGAGTGCAGAGTTCAGCGCTCCAGCTC 1024
|||||
QY 1299 AGCCCAATGAGCAATAACTGTCTGCCAGTGGCAGCGGGAACGGGAGGGGAGTAGCA 1358
|||||
DB 1025 AGCCCAATGAGCAATAACTGTCTGCCAGTGGCAGCGGGAACGGGAGGGGAGTAGCA 1084
|||||
QY 1359 GCTCTGGAGGAGCTCGGGAGGAGTGTCTGTCTGTCTGAGAGCAACGGGCTCTACC 1418
|||||
DB 1085 GCTCTGGAGGAGCTCGGGAGGAGTGTCTGTCTGTCTGAGAGCAACGGGCTCTACC 1144
|||||
QY 1419 CGCTGGCAAAATACAGAAATGCTTCTGGCACTGCGTGAATGGAGTACGTACCAAGCAGA 1478
|||||
DB 1145 CGCTGGCAAAATACAGAAATGCTTCTGGCACTGCGTGAATGGAGTACGTACCAAGCAGA 1204
|||||
QY 1479 ACTGCCAGCGGGCTGTCTGCGACACAGCTGTGATTTGCTGCAACTGGGCATAAACCT 1538
|||||
DB 1205 ACTGCCAGCGGGCTGTCTGCGACACAGCTGTGATTTGCTGCAACTGGGCATAAACCT 1264
|||||
QY 1539 GACCTGGTCTATATTCCTTAGAGTTCAGTCTCTTTTGTCTTAGACATGTTGGCCCTACC 1598
|||||
DB 1265 GACCTGGTCTATATTCCTTAGAGTTCAGTCTCTTTTGTCTTAGACATGTTGGCCCTACC 1324
|||||
QY 1599 TAAAGTCTTCAATAAATCAGCAGTC 1625
|||||
DB 1325 TAAAGTCTTCAATAAATCAGCAGTC 1351
|||||
```

```
RESULT 3
AB025008
LOCUS Homo sapiens TSA1902-L mRNA for novel member of chitinase family,
DEFINITION complete cds.
ACCESSION AB025008
VERSION AB025008.1 GI:6467176
```

AB025008 1354 bp mRNA linear PRI 25-NOV-1999


```

QY 1595 TACCTAAAGTCTCTGCAATAAAA 1616
||||| ||| |||||||||
Db 1493 TACCTAGAGTTCTGCAATAAAA 1514

RESULT 5
AK098814
LOCUS
DEFINITION Homo sapiens cDNA FLJ25948 fis, clone SPM04207, highly similar to
Homo sapiens acidic mammalian chitinase precursor.
ACCESSION AK098814
VERSION AK098814.1 GI:21758941
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens stomach mucosa cDNA to mRNA, clone_lib:STM
clone:STM04207.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Ninomiya,K., Wagatsuma,M., Kanda,K., Kondo,H., Yokoi,T.,
Kodaira,H., Furuya,T., Takahashi,M., Kikkawa,E., Omura,Y., Abe,K.,
Kanihara,K., Katsuta,N., Sato,K., Tanikawa,M., Yamazaki,M.,
Suzuki,Y., Hata,H., Nakagawa,K., Mizuno,S., Morinaga,M.,
Kawamura,M., Sugiyama,T., Irie,R., Otsuki,T., Sato,H.,
Nishikawa,T., Sugiyama,A., Kawakami,B., Nagai,K., Isogai,T. and
Sugano,S.
NEO human cDNA sequencing project
Unpublished
2 (bases 1 to 4250)
Sugano,S. and Suzuki,Y.
Direct Submission
Submitted (08-JUL-2002) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure, Human Genome
Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
(E-mail:cdna@leims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
NEO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction and 5'-end one pass sequencing: Institute of Medical
Science, University of Tokyo, Laboratory of Genome Structure, Human
Genome Center; 3'-end one pass sequencing: RAB; clone selection for
full insert sequencing: RAB and Helix Research Institute.
LOCATION/Qualifiers
1..4250
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="STM04207"
/tissue_type="stomach mucosa"
/clone_lib="STM"
/note="cloning vector: pME18SFL3"
BASE COUNT 1262 a 868 c 966 g 1154 t
ORIGIN

Query Match 64.4%; Score 1046; DB 9; Length 4250;
Best Local Similarity 99.1%; Pred. No. 2.5e-270;
Matches 1052; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 564 TCTTCACTGTCCTGGTGCAGGAATCGTGAAGCTTTTGACGAGGAGCCCAAGCAGATCA 623
||||| ||| |||||||||
Db 3184 TAGTCCTATTATTCTGTAGAAATGCGTGAAGCTTTTGACGAGGAGCCCAAGCAGATCA 3243
||||| ||| |||||||||
QY 624 ACAAGCCAGGCTGATGGTCACTGCTGCAGTAGCTGTGGCATCTCCAATATCCAGTCTG 683
||||| ||| |||||||||
Db 3244 ACAAGCCAGGCTGATGGTCACTGCTGCAGTAGCTGTGGCATCTCCAATATCCAGTCTG 3303
||||| ||| |||||||||
QY 684 GCTATGAGATCCCCCAACTGTCACAGTACCTGGACTACATCCATGTCATGACCTACGACC 743
||||| ||| |||||||||
Db 3304 GCTATGAGATCCCCCAACTGTCACAGTACCTGGACTACATCCATGTCATGACCTACGACC 3363
||||| ||| |||||||||
QY 744 TCCATGCTCCTGGGGGCTACACTGGAGAGACACCCCTCTACAAATACCCGACTG 803
||||| ||| |||||||||

3364 TCCATGGCTCTCTGGAGGGCTACACTGGAGAGAACAGCCCTCTACAAATACCCGACTG 3423
804 ACACGGCAGCAACGCCCTACCTCAATGTGGATTATGTCAATCACTACGAGGACAATG 863
||||| ||| |||||||||
Db 3424 ACACGGCAGCAACGCCCTACCTCAATGTGGATTATGTCAATCACTACGAGGACAATG 3483
||||| ||| |||||||||
QY 864 GAGCACCAGCTGAGAAGCTCATCGTTGGATTCCCTTACCTATGGACACAACTTCATCCCTGA 923
||||| ||| |||||||||
Db 3484 GAGCACCAGCTGAGAAGCTCATCGTTGGATTCCCTTACCTATGGACACAACTTCATCCCTGA 3543
||||| ||| |||||||||
QY 924 GCAACCCCTCAACACACTGGAATTTGGTGGCCCACTCTGGTGTGGTCTCTGCTGGGCCCT 983
||||| ||| |||||||||
Db 3544 GCAACCCCTCAACACACTGGAATTTGGTGGCCCACTCTGGTGTGGTCTCTGCTGGGCCCT 3603
||||| ||| |||||||||
QY 984 ATGCCAAGGACTCTGGGATCTGGGCTTACTACGAGATCTGTACCTTCTCGAAATATGAG 1043
||||| ||| |||||||||
Db 3604 ATGCCAAGGACTCTGGGATCTGGGCTTACTACGAGATCTGTACCTTCTCGAAATATGAG 3663
||||| ||| |||||||||
QY 1044 CCACTCAGGGATGGGATGCCCTCAGGAAGTGCCTTATGCCCTATCAGGGCAATGTGTGGG 1103
||||| ||| |||||||||
Db 3664 CCACTCAGGGATGGGATGCCCTCAGGAAGTGCCTTATGCCCTATCAGGGCAATGTGTGGG 3723
||||| ||| |||||||||
QY 1104 TTGGCTATGACAACATCAAGAGCTTCGATATTAAGGCTCAATGGCTTAAGCACACAAT 1163
||||| ||| |||||||||
Db 3724 TTGGCTATGACAACATCAAGAGCTTCGATATTAAGGCTCAATGGCTTAAGCACACAAT 3783
||||| ||| |||||||||
QY 1164 TTGAGGCGCCATGCTGGGCCATTTGATCTGGATGACTTCACCTGGCACTTTCTGCAACC 1223
||||| ||| |||||||||
Db 3784 TTGAGGCGCCATGCTGGGCCATTTGATCTGGATGACTTCACCTGGCACTTTCTGCAACC 3843
||||| ||| |||||||||
QY 1224 AGGCAAGTTTCCCTTAATCTCCACCTGAAAGAGGCCCTCGGCTCGCAGAGTCAAGTT 1283
||||| ||| |||||||||
Db 3844 AGGCAAGTTTCCCTTAATCTCCACCTGAAAGAGGCCCTCGGCTCGCAGAGTCAAGTT 3903
||||| ||| |||||||||
QY 1284 GCAGGCTCCAGCTCAGGCCATTCAGCCAATACTGTCTCCAGTGGCAGCGGACG 1343
||||| ||| |||||||||
Db 3904 GCAGGCTCCAGCTCAGGCCATTCAGCCAATACTGTCTCCAGTGGCAGCGGACG 3963
||||| ||| |||||||||
QY 1344 GGAGCGGAGTAGCAGCTCTGGAGGAGCTCGGAGGAGCGAGTGTCTGTCTGTCAAG 1403
||||| ||| |||||||||
Db 3964 GGAGCGGAGTAGCAGCTCTGGAGGAGCTCGGAGGAGCGAGTGTCTGTCTGTCAAG 4023
||||| ||| |||||||||
QY 1404 CCAACGGCTCTACCCGCTGGCAATTAACAGAAATGCTTGTGCACTGCTGAATGGAG 1463
||||| ||| |||||||||
Db 4024 CCAACGGCTCTACCCGCTGGCAATTAACAGAAATGCTTGTGCACTGCTGAATGGAG 4083
||||| ||| |||||||||
QY 1464 TCAGCTACCAAGCAACTGCCAGCGGGCTTGTCTTCGACACCACTGTGATTCCTGCA 1523
||||| ||| |||||||||
Db 4084 TCAGCTACCAAGCAACTGCCAGCGGGCTTGTCTTCGACACCACTGTGATTCCTGCA 4143
||||| ||| |||||||||
QY 1524 ACTGGCATAAACCTGACCTGGTCTATATTCCTTAGAGTTCCAGTCTCTTTTGTCTTAGGA 1583
||||| ||| |||||||||
Db 4144 ACTGGCATAAACCTGACCTGGTCTATATTCCTTAGAGTTCCAGTCTCTTTTGTCTTAGGA 4203
||||| ||| |||||||||
QY 1584 CATGTTCCCTACCTAAAGTCTTGCATAAATTAATTCAGCAGTC 1625
||||| ||| |||||||||
Db 4204 CATGTTCCCTACCTAAAGTCTTGCATAAATTAATTCAGCAGTC 4245
||||| ||| |||||||||

RESULT 6
BC034548
LOCUS
DEFINITION Mus musculus, chitinase, acidic, clone MGC:19045 IMAGE:4189080,
mRNA, complete cds.
ACCESSION BC034548
VERSION BC034548.1 GI:21961190
KEYWORDS house mouse.
SOURCE MGC.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1557)
REFERENCE
AUTHORS Strausberg,R.

```

TITLE
JOURNAL
REMARK
COMMENT

Direct Submission
Submitted (24-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amgbcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M., Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAK Plate: 24 Row: h Column: 16
This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.

FEATURES

Source	Location/Qualifiers
	1..1557
	/organism="Mus musculus"
	/db_xref="taxon:10090"
	/map="FVB/N"
	/clone="MGC:19045 IMAGE:4189080"
	/tissue_type="Salivary gland, 10 week old female mouse"
	/clone_lib="NCI CGAP_SG2"
	/lab_host="DH105"
	/note="Vector: pCMV-SPORT6"
	4..1425
CDS	/codon_start=1
	/product="chitinase, acidic"
	/protein_id="AAH34548.1"
	/db_xref="GI:21961191"
	/db_xref="LocusID:81600"
	/translation="MAKLGLTKALLNAQLGSAYNLICYFTNQAQYRPLGSLFKPDDINPCLTHLIYAFAGQMNNEITIEWNDVTLYKAFNDLKRNRSKLKTLLAIGWNFGTAFTPTQSTNSORFTDTSVKFLRGYDFGLDLWEYPGSRGPSPODKHLFTVLVKEEMREAPEQAIENRPLAVAVAGISNIOAGEYIPELSKYLDLHYMTYDLHGSEGYTGENSEPLYKYPIETGSNAVLYNDVYVNKNNGAPAEKLVGPPEGHYTLIRNPSONGICATGGSGAGPYTRQAGFWAYIEICTFLRSRGATEVDASDEVPAVYKANWMLGDYNTKSFVSQAQWLQNKFAGMIWAIDLDDFTGFCPDQGKPLSTLNKAALGISTEITVPVSEPVTTPPGSGSGSSGSGSGSGGSCGCGFCADKADGLYPVADDRNAFWQCINGTVQHQCQAGLVFDTSNCNCNP"
BASE COUNT	410 a 382 c 366 g 399 t
ORIGIN	

Query Match 64.3%; Score 1045.6; DB 10; Length 1557;
Best Local Similarity 83.5%; Pred. No. 3e-270;
Matches 1202; Conservative 0; Mismatches 229; Indels 9; Gaps 1;

QY	101	ACCATCACAAAGCTTATTCTCCTCACAGGTCTTGCTTACTATCTGAATTTGCAGTCGGC	160
Dd	1	ACGATGCCCAAGCTACTTCTCGTCACAGGTCTGGCTCTCTTGTGCTGAATGTCAGTCGGG	60
QY	161	TC TGCCCTACAGCTGACATGCTACTTCCACCACCTGGGCCACAGTACCGCCGACGCTGGG	220
Dd	61	TC TGCCCTACATCTGATATGCTATTTTCACCACTGGGCCACGATCGGCCAGGTCTGGG	120
QY	221	CGCTTCATCGCTGACAACATCGACCCCTCGCTCTGTACCCACCTGATCTACGCGTTTGC	280
Dd	121	AGCTTCAAGCCTGATGACATTAAACCCCTGCCTGTGTACTACCTGATCTATGCTTGTCT	180
QY	281	GGGAGGCAGAACACGAGATCACCCACCTCGAATGAACGATGTGACTCTCTACCAAGT	340
Dd	181	GGGATCGCAACAAATGAGATCACCAACATGGAATGAATGATGACTTCTCTATAAGCT	240

Qy 1421 GTGCAAAATACAGAAATGCTTGTGCACTGGCTGAATGAGTCACTACGACAGCAAC 1480
 Db 1312 GTGCAGATCAGAGAAATCTTTTGGCAGTGCATCAATGAATCAATACACGACCAT 1371
 Qy 1481 TGGCAGCGCGGCTTGTCTGCACACGACGCTGTGATTGTGCAACTGGGCAATACCTGA 1540
 Db 1372 TGTCAGCAGCGGCTTGTCTTGTATACACGCTGTAATGCTGCAACTGGCCATGAACCTAA 1431

RESULT 7

BC011134

LOCUS

DEFINITION

MGC:18771 IMAGE:4165150, mRNA, complete cds.

ACCESSION

BC011134

VERSION

BC011134.1

KEYWORDS

MGC.

SOURCE

house mouse.

ORGANISM

Mus musculus

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

COMMENT

Contact: MGC help desk

Email: egabbs-remail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LLNL)

Sequencing Center

Center code: BCM-HGSC

Web site: http://www.hgsc.bcm.tmc.edu/cdna/

Contact: amg@bcm.tmc.edu

Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,

Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,

Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov

Series: IRAK Plate: 24 Row: e Column: 11.

Location/Qualifiers

1. .1538

/organism="Mus musculus"

/db_xref="taxon:10090"

/map="FVB/N"

/clone="MGC:18771 IMAGE:4165150"

/tissue_type="Salivary gland, 10 week old female mouse"

/clone_lib="NCI-CGAP_SG2"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

323. .1420

/codon_start=1

/product="Similar to eosinophil chemotactic cytokine"

/protein_id="AAH1134.1"

/db_xref="GI:15029822"

/translation="MYSTQNRQTFITSVIKFLRYQYGFGLDLDWEYPSRSGPPQDK

HLFTLVKEMREAFEGEATSNRPLMTAAVAGISNIOAGYEIPELSKYLDIFHW

TYDLHGSWGYTGENSEPLKYPTETGSLNVLNDVYVWYKKNCAPEKLIVGFPEYG

HFLIRNPDSNGIGATSGDGPAGPYTRGAGFWAYEICTFLRSGATEWVNDASQEPY

AYKANWGLNDNKSFSVKAQWLKNQNFGGAMIWAIDLDDFTGSGFCQDKRFPSTLN

KALGISTECPATDPVPEVPTPPGSGSGSGSGSGSGGFCADKADGLYPVADDRN

AFWQCINGITYOQHCQAGLVFTSCNCCNP"

BASE COUNT 393 a 382 c 365 g 398 t

ORIGIN

Query Match 64.1%; Score 1042.2; DB 10; Length 1538;

Best Local Similarity 83.5%; Pred. No. 2.5e-269;

Matches 1198; Conservative 0; Mismatches 228; Indels 9; Gaps 1;
 Qy 106 GACAAAGCTTATCTCTCTCACAGTCTTGTCTTATCTACTGAATTTGAGCTCGGCTCTGC 165
 Db 1 GGCCAAAGCTACTTCTCTGTCTCACAGTCTGGCTCTTCTGCTGAATGCTCAGCTGGGTCTGC 60
 Qy 166 CTACCAAGCTGACATCTTCTACCAACTGGGCGCAGTACGGCCAGGCTCTGGGCGCTT 225
 Db 61 CTACATCTGATATGCTATTTACCAACTGGGCGCAGTATCGGCCAGGCTCTGGGAGCTT 120
 Qy 226 CATGCCCTGACAAATGACACCCCTGCTCTTACCCACTGATCTACGCCCTTTGCTGGGAG 285
 Db 121 CAAGCCTGATGACATTAACCCCTGCTCTGCTACTCACTGATCTATGCTCTGCTGGAT 180
 Qy 286 GCAGAACACGAGATCACCACCATGGAATGGAAGAGTGTGACTCTCTACCAAGCTTCAA 345
 Db 181 GCAGAACATGAGATCACCACCATGGAATGGAATGGAATGGAATGGAATGGAATGGAAT 240
 Qy 346 TGGCCTGAAAATAAGAACACGACGCTGAAACTCTCTGCGCCATTTGGAGGCTGGAACCTT 405
 Db 241 TGACTTGAACACAGGACAGCAAACTGAAACCCCTCTGCAATTTGGAGCTGGAACCTT 300
 Qy 406 CGGAGCTGCCCTTTCTACTGCCATGTTTCTCTCTGAGAACCGCCAGACTTTTCATCAC 465
 Db 301 TGAAGCTGCTCTTCTACTACCATGTTTCTCACTCTCAGAACCGCCAGACTTTTCATCAC 360
 Qy 466 CTGAGTCATCAAAATCTCTGCGCAGTATGATGTTGAGGGCTGAGCTTTGACTGAGGAGTA 525
 Db 361 CTGAGTCATCAAAATCTCTGCGCAGTATGATGTTGAGGGCTGAGCTTTGACTGAGGAGTA 420
 Qy 526 CCCTGGCTCTCTGGGAGCCCTCTCTCAGGACAAAGCATCTCTTCTCTCTCTCTGTCAGGA 585
 Db 421 CCCAGGCTCAGTGGGAGCCCTCTCTCAGGACAAAGCATCTCTTCTCTCTCTCTGTCAGGA 480
 Qy 586 AATGCGTGAAGCTTTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 645
 Db 481 AATGCGTGAAGCTTTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
 Qy 646 TGTGCTGAGTGTCTGGGCTCTCCATATCCAGTCTGCTGCTATGAGATCCCAACATGTC 705
 Db 541 TGTGCTGAGTGTCTGGGCTCTCCATATCCAGTCTGCTGCTATGAGATCCCAACATGTC 600
 Qy 706 ACAGTACCTGGAGTACATCCATGCTACGCTACGCTACGCTACGCTACGCTACGCTACGCT 765
 Db 601 TAAGTACCTGGATTTTCTATCCATGCTATGATGATGATGATGATGATGATGATGATGAT 660
 Qy 766 CACTGGAGAGACAGCCCTCTACAAATACCCGAGTACGACCGGCGAGCAACGCTACCT 825
 Db 661 CACTGGGAGAGTGTCTCTTTACAAATACCCCTACTGAGACTGGTAGCAATGCCTACCT 720
 Qy 826 CAATGTGGATTTATGCTATGAATCTGGAAGAGCAATGGAAGAGCAATGGAAGAGCAATG 885
 Db 721 CAATGTGGATTTATGCTATGAATCTTGGAAAGACATGGAAGAGCAATGGAAGAGCAATG 780
 Qy 886 CGTTGGATTTCCCTACCTATGGACACACCTTCTCTCTGAGCAACCCCTCTCAACACTGGAAT 945
 Db 781 TGTGGATTTCCAGAGTATGGACACACCTTCTCTCTGAGCAACCCCTCTCTGATAATGGAAT 840
 Qy 946 TGTGCCCCCACCCTCTGGTCTGCTGCTGGGCGCTTCTGCTGCGGCGCTTCTGCGGAGTCTG 1005
 Db 841 TGTGCCCCCACCCTCTGGTCTGCTGCTGGGCGCTTCTGCTGCGGCGCTTCTGCGGAGTCTG 900
 Qy 1006 GGTCTTACTAGAGATCTGTACCTTCTCTGAAAATGAGGACCACTCAGGATGGGATGGCC 1065
 Db 901 GGCTTACTATGAGATTTGACCTTTCTGAGAGTGGAGCCACTGAGGCTCTGGATGGCTC 960
 Qy 1066 TCAGGAAGTGCCTTATGCTCTATCAGGCAATGTGTGGTGGCTATGACAACTCAAGAG 1125
 Db 961 CCAAGAAGTGCCTTATGCTCTATTAAGCCCAACGAGTGTGTGGCTATGACAATATCAAG 1020
 Qy 1126 CTTGATATTAAGGCTCAATGGCTTAAGCAACAAATTTGGAGGCGCCATGCTCTGGGC 1185
 Db 1021 CTTGATGTTAAGGCTCAGTGGCTTAAGCAACAAATTTGGAGGCGCCATGCTCTGGGC 1080

```
QY 1186 CATTGATCTGGATGACTTCACTGGCACTTTCTGCAACCAAGGCAAGTTTCCCTCATATCTC 1245
||||| || ||||||||||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Db 1081 CATTGACCTTGATGACTTCACTGGCTTCTTCTGTGATGATGAGGAANAATTTCCCTGACTTC 1140
||||| || ||||||||||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
QY 1246 CACCCTGAAGAGGCCCTCGGCTCGAGAGTCAAGTTCAGCGGCTCCAGCTCAGCCCAT 1305
||||| || ||||||| || || || || || || || || || || || || || || || || ||
Db 1141 TACTTTGAACAAAGCCCTTGGCATATCCACTGAAGTTCACAGCTCCTTGACGTGCTTC 1200
||||| || ||||||| || || || || || || || || || || || || || || || || ||
QY 1306 TCAGGCAATAAAGCTGCTCCAGTGGCAGCGGAGCGGAGCGGAGTACGAGCTCTGG 1365
||||| || ||||||| || || || || || || || || || || || || || || || || ||
Db 1201 CGAGCCAGTGACTCTCTCC-----AGGAAGTGGGAGTGGGGTGGGAAGTCCGG 1251
||||| || ||||||| || || || || || || || || || || || || || || || || ||
QY 1366 AGGAGCTCGGAGGAGTGATTTCTGCTGCTGAGCAACAGCGGCTCTACCCCTGGC 1425
||||| || ||||||| || || || || || || || || || || || || || || || || ||
Db 1252 AGGAAGCTCGGAGGAGTGATTTCTGCTCCGCAAGAGAGATGGCTCTACCCCTGGC 1311
||||| || ||||||| || || || || || || || || || || || || || || || || ||
QY 1426 AATAACAGAAATGCTTCTGCACTGCTGATGAATGAGTACAGTACAGCAGAACTGCCA 1485
||||| || ||||||| || || || || || || || || || || || || || || || || ||
Db 1312 AGATGACAGAAATGCTTTTGGCAGTGCAATGAATGCAATACAGCAGCAGATTGTCA 1371
||||| || ||||||| || || || || || || || || || || || || || || || || ||
QY 1486 GCCCGGCTGTCTTCGACACCACTGATGCTGCTCAACTGGGCATAAACCTGA 1540
||||| || ||||||| || || || || || || || || || || || || || || || || ||
Db 1372 AGCAGGCTGTGTTTGTATACCACTGTAATTGCTGCAACTGGCCATGAACCTAA 1426
||||| || ||||||| || || || || || || || || || || || || || || || || ||

RESULT 8
AF154571
LOCUS AF154571 1526 bp mRNA linear ROD 08-FEB-2000
DEFINITION Mus musculus putative chitinase precursor (YNL) mRNA, partial cds.
ACCESSION AF154571
VERSION AF154571.1 GI:6934189
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1526)
AUTHORS Price,P.A., Harris,S.C. and Williamson,M.K.
TITLE YNL, A Putative Mouse Chitinase
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1526)
AUTHORS Price,P.A., Harris,S.C. and Williamson,M.K.
TITLE Direct Submission
JOURNAL Submitted (26-MAY-1999) Biology, University of California, San
Diego, 9500 Gilman Drive, Mail Code 0368, La Jolla, CA 92093-0368,
USA
FEATURES
source
location/Qualifiers
1..1526
/organism="Mus musculus"
/strain="ATCC147668"
/db_xref="ATCC:147668"
/db_xref="taxon:10090"
/clone="IMAGE 1209227"
/tissue_type="skin"
<1..1526
/gene="YNL"
<1..1419
/gene="YNL"
/codon_start=1
/product="putative chitinase precursor"
/protein_id="AAF31644.1"
/db_xref="GI:6934190"
/translaton="AKLLIVTGLALLNAQLGSAYNLICFTNWAQYRPLGSKFPPDD
INPLCTHLIYAFACQMNNEITIEHNDVLYKAENDLKNRSKLTLILAIGWNEFG
APFTMVSTQNRQFITSVKFLQYDGLDLDWEYPGSRSGPPDKHLFTLVKYE
MREAFQEAIENRRLPMVAAGVSIQAGYEIPELSKYIDFIHVTYDILHGSWE
GYTGNSPLYKYPTTGSNAYLVNDVVMNKNNGAPAEKLIIVGPEYHTFILRNP
DNIGAPTSGDGPAGPYRQAGFWAYEICTFLRSQATEVWDASOEVPYAYRANWL
YNIKFSVKAWLQNNFGGAMINWIDLDDETSFCDOGKEFLPLSTLNKALISTEG
CTAPDVPSEPTVTPGSGSGSGSGSGSGGFCADKADGLIPVADDRNFWQCLINGI
TYQHCQAGLVDFDSCNCCNWP"
<1..60
sig_peptide
```

```
mat_peptide /gene="YNL"
61..1416
/misc_feature /gene="YNL"
394..417
/misc_feature /product="putative chitinase"
415..417
/misc_feature /gene="YNL"
415..417
/misc_feature /note="putative conserved region; active site"
415..417
/misc_feature /gene="YNL"
415..417
/misc_feature /note="putative invariant Glu; active site"
415..417
BASE COUNT 383 a 381 c 364 g 398 t
ORIGIN
Query Match 64.1%; Score 1041.8; DB 10; Length 1526;
Best Local Similarity 83.5%; Pred. No. 3.2e-269;
Matches 1197; Conservative 0; Mismatches 227; Indels 9; Gaps 1;
QY 108 CAAAGCTTATTTCTCTCACAGGTCTTGTCTTATCTGAATTTGCAGCTCGGCTCTGCGCT 167
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Db 2 CCAAGCTACTTCTCGTCAAGGTCTGGCTCTTCTGCTGAATGCTCAGCTGGGCTCTGCGCT 61
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
QY 168 ACCAGCTGACATGCTACTTTCACCAACTGGGCCCCAGTACGGCCAGGCTGGGCGGCTTCA 227
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Db 62 ACAATCTGATATGCTATTTTCCCAACTGGGCCCCAGTATCGCCAGGCTCTGGGAGGCTTCA 121
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
QY 228 TGCCTGACAACTCGACCCCTGCTCTGTACCCACCTGATCTACGCTTTTGTGGGAGGC 287
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Db 122 AGCTGATGACATTAACCCCTGCTGTACTACCTGATCTATGCTTTGCTGGGATGC 181
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
QY 288 AGAACACAGAGTACACCACTGGAATGGAACGATGTGACTCTCTACCAAGCTTTCAATG 347
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Db 182 AGAACAAATGAGATCACCACTGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGA 241
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
QY 348 GCCTGAAAATTAAGAACAGCAGCAGCTGAAAACCTCTCTGCGCATTTGAGGCTGAACTTCG 407
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Db 242 ACTTGAANAACAGGAACAGCAAACTGAAAACCCCTCTGCGCAATTTGAGGCTGAACTTCG 301
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
QY 408 GACTGCCCCCTTTTCACTGCGCATGTTTCTACTCTCTGAGAACCCGACAGCTTTTCATCACCT 467
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Db 302 GACTGCTCTCTTCACTACCATGTTTCCACTTCTCAGAACCCGACAGCTTTTCATCACCT 361
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
QY 468 CAGTCATCAAAATTCCTGCGCCAGTATGATTTGACGGGCTGGAGCTTTGAGTGGAGTACC 527
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Db 362 CAGTCATCAAAATTCCTGCGCTCAGTATGGGTTTGTGAGCTGGAGCTGGAGTACC 421
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
QY 528 CTGGCTCTGCTGGGAGCCCTCTCAGACAGCATCTTCACTGCTGCTGCTGCTGCTGCTGCT 587
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Db 422 CAGGCTCAGCTGGGAGCCCTCTCTCAGGACAAGCATCTTCTACTGCTGCTGCTGCTGCTGCT 481
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
QY 588 TCGTGAAGCTTTTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 647
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Db 482 TCGTGAAGCTTTTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 541
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
QY 648 CTGCACTAGCTGCTGGCATCTCCAATATCAAGTCTGGCTATGAGATCCCCCACTGTAC 707
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Db 542 CTGCTAGCTGCTGGGAGTTTCCAACATCCAGCTGGCTATGAGATCCCCCACTGTAC 601
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
QY 708 AGTACCTGCACTACATCCATGTCATGACCTTACACCTTCCATGCTCTCTGGAGGCTTACA 767
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Db 602 AGTACCTGCACTACATCCATGTCATGACATATGACCTTCCATGCTCTCTGGAGGCTTACA 661
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
QY 768 CTGGAGAGAACAGCCCTCTTACAAATACCCGACTGACACCCGAGCAGCAGCCTCACTCA 827
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Db 662 CTGGAGAGAACAGCTCTCTTACAAATACCCGACTGACATGCTGAGTACCTCACTCA 721
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
QY 828 ATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 887
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Db 722 ATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 781
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
QY 888 TTGGATTTCCCTACCTATGACACAACTTCTCTGAGCAACCCCTCCACACTGGGAATTG 947
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Db 782 TTGGATTTCCAGAGTATGACACAACTTCTCTGAGCAACCCCTCCACACTGGGAATTG 841
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
```

```
Qy 948 GTGCCCCACCTCTGGTGGTCTCTCTGGGCGCTATGCCAAGAGCTCTGGGATCTGGG 1007
Db 842 GTGCCCCACCTCTGGTGGTCTCTCTGGGCGCTATGCCAAGAGCTCTGGGATCTGGG 901
Qy 1008 CTTACTACGAGACTGTACCTTCTTGAAAATGGAGCCACTCAGGAGTGGATGCCCTC 1067
Db 902 CCTACTAGATTTGACCTTCTTGAAAATGGAGCCACTCAGGAGTGGATGCCCTC 961
Qy 1068 AGAAGTGCCTTATGCTGATCAGGCAATGTGTGGTGGTATGCAACATCAAGAGCT 1127
Db 962 AGAAGTGCCTTATGCTGATCAGGCAATGTGTGGTGGTATGCAACATCAAGAGCT 1021
Qy 1128 TCGATATTAAGGCTCAATGCTTAAAGCAACAAATTTGAGGCGCCATGCTTGGGCCA 1187
Db 1022 TCAGTGTAAAGGCTCAATGCTTAAAGCAACAAATTTGAGGCGCCATGCTTGGGCCA 1081
Qy 1188 TTGATCTGGATGACTTCACTGGCATTCTTCGACACAGGCAAGTTTCCCTATCTCA 1247
Db 1082 TTGACCTGTGATGACTTCACTGGCATTCTTCGACACAGGCAAGTTTCCCTATCTCA 1141
Qy 1248 CCGTGAAGAAGGCGCTCGGCTGCGAGTGCAGAGTGCAGGCTCCAGCTCAGCCATTG 1307
Db 1142 CTTTGACAAAGCCCTGGCATATCACTGAAGTTCGACAGCTCTGAGCTGCCCTCG 1201
Qy 1308 AGCAATTAAGTCTGCTCCAGTGGCAGCGGAGCAAGCGGAGTGGAGTCTGGAG 1367
Db 1202 AGCAGTGAATCTCTCTCC-----AGGAAGTGGGAGTGGGAGTCCGGAG 1252
Qy 1368 GCAGCTGGGAGGAGTGGATTCTGTGCTGTCAGAGCAACGCGCTTACCCGCTGCA 1427
Db 1252 GAGCTCTGGAGGAGTGGATTCTGTGCTGTCAGAGCAACGCGCTTACCCGCTGCA 1312
Qy 1428 ATACAGAAATGCTTCTGCACTGCGTGAATGAGTGCAGTACCGAGCAAGTGCAGG 1487
Db 1312 ATGACAGAAATGCTTCTGCACTGCGTGAATGAGTGCAGTACCGAGCAAGTGCAGG 1372
Qy 1488 CCGGCTGTCTTCGACACAGTGTGATTGCTGCACTGGGATCAACCTGA 1540
Db 1372 CAGGCTGTCTTCGACACAGTGTGATTGCTGCACTGGGATCAACCTGA 1425

RESULT 9
AF290003
LOCUS
DEFINITION Mus musculus acidic mammalian chitinase precursor, mRNA, complete
cgs.
ACCESSION AF290003
VERSION AF290003.1
KEYWORDS GI:12597290
SOURCE
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1530)
Boot.R.G., Blommar, E.F., Swart, E., Ghaubari-van der Vlugt, K.,
Blj.N., Moe, C., Place, A. and Aerts, J.M.
Identification of a novel acidic mammalian chitinase distinct from
chitotriosidase
J. Biol. Chem. 276 (9), 6770-6778 (2001)
JOURNAL
MEDLINE 21125893
PUBMED 11085997
REFERENCE 2 (bases 1 to 1530)
AUTHORS Boot.R.G., Verhoeck, M., Swart, E. and Aerts, J.M.
TITLE Direct Submission
JOURNAL Submitted (26-JUL-2000) Dept. of Biochemistry, Academic Medical
Center, University of Amsterdam, Meibergdreef 15, Amsterdam 1105
AZ, The Netherlands
FEATURES
source
1. .1530
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
/sex="male"
```

CDS

```
/tissue_type="lung"
/dev_stage="9-11 weeks"
3. .1424
/EC_number="3.2.1.14"
/note="chitinase precursor; AMCase"
/codon_start=1
/product="acidic mammalian chitinase precursor"
/protein_id="AAG60018.1"
/db_xref="GI:12597291"
/translation="MAKLLVLTGLALLINAOGLSAYNLICFTNWAQYRPLGSGFKPD
DINPCLTHLIYAFAGQNEITIEWDLVLYKAFNDLKNRNSKLLTLLAIGWNFG
TAPFTVMTSONRQFTITSVKFLROYDGLDLMWEPGSRGSPDQKHLFTVLVK
EMREAEQEAIESNRLPMVTAAGVGSINQIAGYIPELSKYLDTIHVYDLHGSW
EGYTGNSPLKYPTETGSNAYLVNDVVMYKNGNAPAEKLVGPPEYGHFTFLRP
SDNGIGAPTSDDGPAGATROAGFWAYEICTFLRSGATEVWDASOEPVYKANEWL
GYDNIKSFYKAQWLKONFNGGAMWAILDLDFTSGCDGKFLPLSTLNKALGISTE
GCTAPDVPSEPTTPPGSGSGSGSGSGSGSGGFCADKADGLIPVADDRNAPWOCING
ITVQHQHAGLGVDFDTSNCNWP"
BASE COUNT 382 a 383 c 366 g 399 t
ORIGIN
Query Match 64.1%; Score 1041.4; DB 10; Length 1530;
Best Local Similarity 83.3%; Pred. No. 4.1e-269;
Matches 1199; Conservative 0; Mismatches 231; Indels 9; Gaps 1;
Qy 102 CCATGACAAAGCTATTCTCTCACAGGCTTGTCTCTTATCTATCTAGCTCGCT 161
Db 1 CGATGCCAAGCTACTTCTCTCACAGGCTTGTCTCTTCTCTGCTGCTGCTGCTG 60
Qy 162 CTGCTTACAGCTGACATGCTTCTTACCAACTGGGCGCCAGTACCGGCCAGGCTGGGC 221
Db 61 CTGCTTACAACTGATGCTTATTCACCAACTGGGCGCCAGTATCGGCCAGGCTGGGGA 120
Qy 222 GCTTCATGCTGACACATCGACCCCTGCTCTGTACCCACCTGTACGCTCTGCTG 281
Db 121 GCTTCAGGCTGATGACATTAACCCCTGCTGTACTCACCTGATCTATGCTCTGCTG 180
Qy 282 GGAGGAGAGAACACGAGATCACCACTCAATGGAACGATGCTGCTCTTACCAAGCTT 341
Db 181 GGATGACAGAACATGATGATCACCACTCAATGGAACGATGCTGCTCTTATTAAGCTT 240
Qy 342 TCAATGGCTGAAATAAGAACAGCAGCTGAAACTCTCTGCGCCATTTGAGCTGGG 401
Db 241 TCAATGACTTGAATAAGAACAGCAGCAAACTGAAACCTCTCTGCGCCATTTGAGCTGGG 300
Qy 402 ACTTGGGAGCTGCGCTTTTTCACGCTGCTGCTTCTACTCTGAGAACCGCAGACTTCA 461
Db 301 ACTTGGAGCTGCTCTCTTCTCCTACCATGCTTCTCCTGAGAACCGCAGACTTCA 360
Qy 462 TCACCTCAGTCAATAATCTCTGCGCCAGTATGAGTTTACGGGCTGGACTTTGACTGGG 521
Db 361 TTACCTCAGTCAATAATCTCTGCGTCAAGTATGGTTTGTGAGCTGGACTGGG 420
Qy 522 AGTACCTGCTCTCTGTTGGAGCCCTCTCTCAGAACAGCATCTCTTCACTGCTCTGGTGC 581
Db 421 AATACCCAGGCTCACCTGGGAGCCCTCTCAGAACAGCATCTCTTCACTGCTCTGGTGA 480
Qy 582 AGAAATGCTGAGCTTTTGGAGCAGGCGCAAGCAGATCAACAAGCCAGCTGATGG 641
Db 481 AGAAATGCTGAGCTTTTGGAGCAGGAGCTATTGAGACAGAGCTTATGAGACAGCTGATGG 540
Qy 642 TCACCTGCTGAGTGTGCTGCTGCTTCCAAATATCCAGTCTGGCTATGAGATCCCCAAC 701
Db 541 TTACTGCTGCTGAGTGTGCTGCTTCCAAATATCCAGCTGCTGCTATGAGATCCCCAAC 600
Qy 702 TGTCAGTACCTGGAGTACATCATGATGATGATGATGATGATGATGATGATGATGATG 761
Db 601 TTTCTAAGTACCTGGATTTTCTATGATGATGATGATGATGATGATGATGATGATGATG 660
Qy 762 GCTTACACTGGAGAGAACAGCCCTCTTACAAATATCCAGCTGACACCGCAGCAAGCTCT 821
Db 661 GCTTACACTGGAGAGAAATAGTCTCTTACAAATATCCAGCTGATGAGCTGATGAGCTCT 720
```

BASE COUNT	280 a	337 c	309 g	262 t
ORIGIN	IPVANKRMRWCVNGVYIQQNCQAGLVFDFISCDCCNWA			

Query Match 63.8%; Score 1037.2; DB 9; Length 1188;
Best Local Similarity 99.7%; Pred. No. 5.4e-268;

Qy 584 GAAATGCGTGAAGCTTTTGTGACGAGGGCCAAGCAGATCAACAAAGCCCAGGCTGATGGTC 643

Db 147 GAAATGCGTGAAGCTTTTGAGCAGGAGGCCAAGCAGATCAACAAGCCAGGCTGATGGTC 206

[illegible]

QY 704 TCACAGTACCTGGACTACATCCATGTATGACCTACGACCTCCATGGCTCCTGGGAGGGC 763

Db 267 TCACAGTACCTGGACACATCCATGTCATGACCTACGACCCTCCATGGCTCCTGGAGGGC 326

[illegible]

QY 824 CTCGAATGTGGATTATGTCATGAAC TACTGGAAGGACAAATGGACCACCAGCTGAGAGACCTC 883

D_b 387 CTCAATGTGGATTATGTCATGAAC TACTGGAAGGACAATGGAGCACCAGCTGAGAGGTC 446

QY
884 ATCGTGGATTCCCTACCTATGGACACAACTTCATCTCGAGCAACCCCTCCCAACACTGGA 943

944 ATTGGTGGCCCCACCTTCGTGGTGCTGGCCCATAGGCCAAGGAGTCTGGGCATC

Db 507 ATTGGTCCCCCACCCTCTGGTCTGGTCCCTGCTGGGCCCTATGCCAAGGAGTCTGGGATC 566

QY 1004 TGGSCITACTACGAGATCTGTACCTTCTTGAAAAATGGAGCCACTCAGGGATGGGATGCC 1063

1054 CCTCAGGAAGTGCCTATGCCCTATCAGGGGAAATGTGTGGGTTGGCTATGACACATCAAG 1123

Db 627 CCTCAGGAAGTGCCCTTATGCCCTATCAGGGCAATGTGTGGGTGGCTATGACAAAGTCAAG 686

QY 1124 AGCTTCGATATTAAAGGCTCAATGGCTTAAGCACACAATAATTTGGAGGGCCCATGGTCTGG 1183

[illegible]

Db 747 GCATTGATCTGGATGACTTCACCTGGCACTTTCTGCAACCAAGGCAAGTTTCCCCTAATC 806

Qy 1244 TCCACCTGAAAGAGGGCCCTCGGCCTGCAGAGTGAAGTTGCACGGCTCCAGCTCAGCCC 1303

DD 807 TCCACCCCTGAAAGAAAGGCCCTTGGCCCTGGCAGAGTGCAGAGTTGCACGGCTCCAGGCTCAGGCC 808


```
QY 1304 ATTGAGCCAAATACTGCTGCTCCAGTGGCAGCGGGAACGGAGCGGAGTAGCAGCTCT 1363
|||||
Db 867 ATTGAGCCAAATACTGCTGCTCCAGTGGCAGCGGGAACGGAGCGGAGTAGCAGCTCT 926
|||||
QY 1364 GGAGCGAGCTCGGAGGAGTGGATTCTGCTGTCAGAGCCAGCGGCTTACCCCGTG 1423
|||||
Db 927 GGAGCGAGCTCGGAGGAGTGGATTCTGCTGTCAGAGCCAGCGGCTTACCCCGTG 986
|||||
QY 1424 GCAATAACAGAAATGCTTCTGCGACTGCGTGAATGGAGTCAGTACCAGCAAGCTGC 1483
|||||
Db 987 GCAATAACAGAAATGCTTCTGCGACTGCGTGAATGGAGTCAGTACCAGCAAGCTGC 1046
|||||
QY 1484 CAGCCCGGGCTGCTTCGACACCAAGCTGTGATTGCTGCAACTGGGCAATAACTGACCT 1543
|||||
Db 1047 CAGCCCGGGCTGCTTCGACACCAAGCTGTGATTGCTGCAACTGGGCAATAACTGACCT 1106
|||||
QY 1544 GGTCTATATCCCTAGAGTCCAGTCTCTTTGCTTAGGACATGTTGCCCTACTAAAG 1603
|||||
Db 1107 GGTCTATATCCCTAGAGTCCAGTCTCTTTGCTTAGGACATGTTGCCCTACTAAAG 1166
|||||
QY 1604 TCCTGCAATAAAATCAGCAGTC 1625
|||||
Db 1167 TCCTGCAATAAAATCAGCAGTC 1188
|||||
```

RESULT 11

```
LOCUS M94584 1525 bp mRNA linear ROD 15-JUN-2001
DEFINITION Mus musculus secretory protein precursor (Yml) mRNA, complete cds.
ACCESSION M94584
VERSION M94584.2 GI:11140876
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1525)
Chang,N.-C.A., Liu,C.-H. and Chang,A.C.
Molecular characterization of a secretory protein (Ym-1)
transiently expressed by activated murine peritoneal macrophages
Unpublished
2 (bases 1 to 1525)
Chang,N.C., Hung,S.I., Hwa,K.Y., Kato,I., Chen,J.E., Liu,C.H. and
Chang,A.C.
A macrophage protein, Yml, transiently expressed during
inflammation is a novel mammalian lectin
J. Biol. Chem. 276 (20), 17497-17506 (2001)
21264517
11297523
3 (bases 1 to 1525)
Chang,N.-C.A., Liu,C.-H. and Chang,A.C.
Direct Submission
Submitted (27-APR-1993) Microbiology and Immunology, National
Yang-Ming University, 155, Sect. II, Li-Noon St, Taipei, Taiwan
112, R.O.C.
4 (bases 1 to 1525)
Chang,N.-C.A., Liu,C.-H. and Chang,A.C.
Direct Submission
Submitted (13-NOV-2000) Microbiology and Immunology, National
Yang-Ming University, 155, Sect. II, Li-Noon St, Taipei, Taiwan
112, R.O.C.
REMARK
Sequence update by submitter
COMMENT On Nov 13, 2000 this sequence version replaced gi:202441.
FEATURES
source 1..1525
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_lib="PEC in lambda gtl1 and ZAPII"
/dev_stage="adult"
1..1525
/gene="Yml"
13..1209
gene
CDS
```

```
/gene="Yml"
/codon_start=1
/product="secretory protein precursor"
/protein_id="AAB62394.2"
/db_xref="GI:11140877"
/translation="MAKLILVTGLAILNLVOLGSSYOLMCYIYTSWAKDRPIEGSFKPG
NIDPCJCTHLIYAFAGQNEITYTHQDLRDYALNGLKDKNTELRKLLAIGWKFV
PAPFSAMVSPQNRQIFIQSVIRFQYDFNLDWQYPGSRGSPKDRHLFSLVK
EMKAPESVEKDIPIRLLTSGAGIIDIYKSGYKIPELSQSLDYIQVMTYDLHDPK
DGYTGNSPLYKSPYDIDKSLNVDISIYWKDHGAASEKLIVGFPAYGHTFTLSDP
SKYIGAPTISTGPPGKYTDESGLLAYEYVCTFELNEGATEVWDAPQEVFAYQGNWV
GYDNVRSFKLKAOWLKDNLNGAVVWPLDMDDFSGFCHORHFFLTSTLKGDLNIHSA
SCKGPY"
13..75
/gene="Yml"
76..1206
mat_peptide
/gene="Yml"
/product="secretory protein"
1487..1492
/polyA_signal
/gene="Yml"
1507
polyA_site
/gene="Yml"
343 c 335 g 413 t
BASE COUNT 434 a 343 c 335 g 413 t
ORIGIN
Query Match 43.2%; Score 701.8; DB 10; Length 1525;
Best Local Similarity 74.3%; Pred. No. 1.2e-177;
Matches 886; Conservative 0; Mismatches 307; Indels 0; Gaps 0;
QY 101 ACCATGACAAAGCTTATTCCTCCTCACAGGCTGCTCTTATCTATCTGAAATTTGCAGCTCGG 160
|||||
Db 10 ACCATGACAAAGCTTATTCCTCCTCACAGGCTGCTCTTATCTGAAATTTGCAGCTCGG 69
|||||
QY 161 TCTGCTTACCAGCTGACATGCTACTTCCACCACTGGCCGAGTACCGCCAGGCGCTGGG 220
|||||
Db 70 TCTTCTTACCAGCTGATGCTGCTACTATACCAGTTGGGCTAAGCAGCGCAATAGAAGG 129
|||||
QY 221 CGCTTCACTGCTGACCAATCGACCCCTGCTCTGTACCCACCTGATCTACGCTTTGCT 280
|||||
Db 130 AGTTTCAAACTGCTGATATTTGACCCCTGCTGTGTACTCACCTGATCTATGCTTTGCT 189
|||||
QY 281 GGAGGCGAGAACACAGAGATCACCACTCGAATGGAAACGATGTGACTCTCTACCAAGCT 340
|||||
Db 190 GGAATGCGAATATGAGATCACTATACACATGAGCAAGACTTGCCTGACTATGAAGCA 249
|||||
QY 341 TTCAATGGCTGAAAAATAAGAACAGCAGCTGAAAACTCTCTGGCCATTGGAGGCTGG 400
|||||
Db 250 TTGAATGGCTGAAAGACAGACACTGAGCTAAAAACTCTCTGGCCATTGGAGGATGG 309
|||||
QY 401 AACTTGGGACTGCCCTTTTCACTGCCATGGTTTCTACTCTCTGAGAACGCCGACATTC 460
|||||
Db 310 AAGTTTGGACTGCCCGTTTCACTGCCATGGTTTCTACTCTCTCAGAACCCGTCAGATATTC 369
|||||
QY 461 ATCACTCAGTCATCAAAATTCCTGCGCCAGTATGATTTTACGGGCTGGACTTTGACTGG 520
|||||
Db 370 ATTCACTCAGTCATCAAAATTCCTGCGCCAGTATGATTTTACGGGCTGGACTTTGACTGG 429
|||||
QY 521 GAGTACCTGCTCTCTGCTGGAGCGCTTCTCAGGACAGCATCTCTTCACTGTCTCTGGT 580
|||||
Db 430 CAGTACCTGCTCTCTGCGAGGAGCGCTTCTTAAAGGACAAACATCTCTTCACTGTCTGGT 489
|||||
QY 581 CAGGAATGCGTGAAGCTTTTGAAGAGGAGGCGCAAGCAGATCAACAGCCAGCGGATGATG 640
|||||
Db 490 AAGGAATGCGTGAAGCTTTTGAAGAGGAGGAGTCTGTGGAGAAAGACATCTCAAGGCTGCTA 549
|||||
QY 641 GTCACTGCTCAGTGTGCTGGCATCTCCAATATCCAGTCTGGCTATGATAGATCCCCCAA 700
|||||
Db 550 CTCACCTCCACAGGAGCGAGGAATCAATTGAGCTAAGTCTGGGTCAAGATCCCTCGAA 609
|||||
QY 701 CTGTCACAGTACCTGGACTACATCCATGTGATGACCTTACGACCTCCATGCTCTCTGGGAG 760
|||||
Db 610 CTGCTCAGTCTCTTGACTATATTCAGGTATGATGATCTCCATGATCTCCATGATCTTANGAT 669
|||||
```



```
QY 761 GGCTACACTGGAGAACAGCCCTCTACAAATACCGACGACACCGCGACAGCC 820
Db 670 GGCTACACTGGAGAAATAGTCCCTCTATAATCCATATCCATATGGAAGAGTGT 729
QY 821 TACCTCAATGTGGATATGCTGAACACTGAGGAGCAATGAGGACCACTGAGAG 880
Db 730 GATCTCAATGTGGATATGCTGAACACTGAGGAGCAATGAGGACCACTGAGAG 789
QY 881 CTCATCGTGGATTCCTACCTATGACACAACTTCATCCTGAGCAACCCCTCAAC 940
Db 790 CTCATCGTGGATTCCTACCTATGAGGACATGAGGACCACTTCATCCTGAGT 849
QY 941 GGAATGTGGCCCCCACCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1000
Db 850 GGAATGTGGCCCCCACCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 909
QY 1001 ATCTGGGCTTACTACGAGATCTGATACCTTCTGAAATGAGGACCACTGAGG 1060
Db 910 CTCCTGGCTTACTACGAGATCTGATACCTTCTGAAATGAGGACCACTGAGG 969
QY 1061 GCCCTCAGGAAGTGCCTTATGCTATCAGGCAATGTGTGGTGGCTATGACA 1120
Db 970 GCCCTCAGGAAGTGCCTTATGCTATCAGGCAATGTGTGGTGGCTATGACA 1029
QY 1121 AAGAGCTTCGATTAATGAGCTCAATGGCTTAAGCAACAAATTTGAGGCGCC 1180
Db 1030 AGGAGCTTCAAGTTGAAGCTCAGTGGCTCAAGGCAACAAATTTAGGAG 1089
QY 1181 TGGGCAATGATCTGATGATCTGATGATCTGATGATCTGATGATCTGATG 1240
Db 1090 TGGGCAATGATCTGATGATCTGATGATCTGATGATCTGATGATCTGATG 1149
QY 1241 ATCTCCACCCTCAAGGAGCCCTCGGCTGAGAGTCAAGTTGACGCGCTCC 1293
Db 1150 ACATCTACTTAAAGGAGATCTCAATATACAGTCAAGTTGCAAGGCCCC 1202

RESULT 12
D87757
LOCUS D87757 1506 bp mRNA linear ROD 12-JAN-2000
DEFINITION Mus musculus mRNA for ECF-L precursor, complete cds.
ACCESSION D87757.2
VERSION 2
KEYWORDS ECF-L precursor.
SOURCE Mus musculus bone marrow cdna to mRNA.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Ohashi.M., Arita.H. and Hayai.N.
Identification of a novel eosinophil chemotactic cytokine (ECF-L)
as a chitinase family protein
J. Biol. Chem. 275 (2), 1279-1286 (2000)
JOURNAL J. Biol. Chem. 275 (2), 1279-1286 (2000)
MEDLINE 20092902
REFERENCE 2 (bases 1 to 1506)
AUTHORS Ohashi.M.
TITLE Direct Submission
JOURNAL Submitted (06-SEP-1996) Makoto Ohashi, The University of
Tokushima, Faculty of Integrated Arts and Sciences; 1-1
Minami-Johsanjima, Tokushima, Tokushima 770, Japan
(E-mail:ohashieias.tokushima-u.ac.jp, Tel:0886-56-7261,
Fax:0886-56-7298)
COMMENT On Oct 7, 1999 this sequence version replaced gi:1545818.
Sequence updated (05-Oct-1999).
FEATURES
source Location/Qualifiers
1..1506
/organism="Mus musculus"
/db_xref="taxon:10090"
/tissue_type="bone marrow"
1..1506
/gene="mECF-L"
1..1197
/gene="mECF-L"
```

```
/codon_start=1
/product="ECF-L precursor"
/protein_id="BAAL3458.2"
/db_xref="GI:6015436"
/translation="MAKLILVYGLAILLVQLSSVQLMCIYTSWAKDRPIBGSFPRFG
NDPCLCTHLIYAFAGMNEIYTHEQDLRDY EALNGUKDNKTELUKTLUAIGGMRFG
PASFAWSTVSTPQNRIFIQSVIRFLQYNFDNLNDMWPYSGSPDKHFLSVLVK
EMRKAFEESEKDIPLRLTSTAGIIDIIVKSGYKIPELSGLDYIOYMTYDLHDPK
DYGTEGNSPLYKSPYDYGKADNLVDSIISYKWDHGAASEKLVGPGVCHTFILSDP
SKTGICAPTISTGPBPCKYDSESLIAYEVCTFLNAGATEVMDADPEVPYAVQGNWV
GYDNVRSFKLKAQWLKDNLNGAVVWPLMDDFSGSFCHQRHFLPTSLTKGLDNLHSA
SCKGPI"
BASE COUNT 423 a 338 c 333 g 412 t
Query Match 42.9%; Score 697.2; DB 10; Length 1506;
Best Local Similarity 74.1%; Pred. No. 2e-176;
Matches 882; Conservative 0; Mismatches 308; Indels 0; Gaps 0;
QY 104 ATGACAAAGCTTATCTCCTCACAGCTTGTGTCTTATATCTGAGCTTCGCTCT 163
Db 1 ATGGCCAAGCTCATCTTCTGACAGCTGGCAATCTTCTGAACGCTACAGCTGGGATCT 60
QY 164 GCTACACGCTGACATGCTACTTACCAACTGGGCCCCAGTACCGGCCAGCCCTGGGCGC 223
Db 61 TCCTACAGCTGATGTCTACTATACCACTGGGCTTAAGGACAGGCCAATAGAGGAGT 120
QY 224 TTCATGCTTACCAACATCGACCCCTGCTGTACCCAGCTGATCTACCGCTTGTCTGGG 283
Db 121 TTCAAACCTGGTAAATATGACCCCTGCTGTGTACTGACCTGATCTATGCTTGTGGGA 180
QY 284 AGCGAAGCAACGAGATCACCACTCGAATGGAAGATGTGTACTCTTACCAAGCTTTC 343
Db 181 ATGCAGAAATAGAGATCACTTACACACATGACCAAGACTTGGCTGACTATGAAGCAT 240
QY 344 AATGGCTGAAATTAAGAACAGCCAGCTGAAATCTCTCTGGCCATTTGAGGCTGGAAC 403
Db 241 AATGGTCTGAAGACAGAACACTGAGCTAAATCTCTCTGGCCATTTGAGGATGGAAG 300
QY 404 TTCGGAGCTGCCCTTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 463
Db 301 TTTGGACCTGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 464 ACCTCAGTCATCAATTCCTGCGCCAGTATGATGATGATGATGATGATGATGATGATG 523
Db 361 CAGTCAGTATCAGATTCCTCTGCTCAATATCACTTTGATGCTCAACCTGAGCTGGCAG 420
QY 524 TACCTCGGCTGCTGCTGGGAGCCCTCTCTCAGGACCAAGATCTCTTCTACTGCTGCTG 583
Db 421 TACCTCGGCTGCTGAGGAGCCCTCTCTTAAGGACCAACATCTCTTCTAGTGTCTCTG 480
QY 584 GAATTCGCTCAAGCTTTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 643
Db 481 GAAATCGTAAAGCTTTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540
QY 644 ACTGCTGAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 703
Db 541 ACTTCCACAGGACGAGGAATCATTTGAGCTAATCAAGTCTGGGTACAAGATCCCTGAG 600
QY 704 TCACGATACCTGGACTACATCCATGATGATGATGATGATGATGATGATGATGATGAT 763
Db 601 TCTCAGCTCTTCTGCTATATTTAGGTCATGACATATGATCTCCATGATCTCAAGGATGG 660
QY 764 TACACTGGAGAGAACAGCCCTCTTACAAATACCCGACTGACACCGGACGACAGCCCT 823
Db 661 TACACTGGAGAGAAATATGCTCCCTCTTAAATATCCCATATGACATTTGAAAGAGTGT 720
QY 824 CTCATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 883
Db 721 CTCATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
QY 884 ATCGTTGGATTCCTTACCTATGAGACAACTTCTCTGAGCAACCCCTCCACACTGGA 943
```

Db 781 ATTCTGGGATTTCCAGCATATATGGGCGATACCTTTATCTGAGTGCACCTTTCAAGACTGGA 840

Qy 944 ATTGGTCCCCCACTCTGGTCTGGTCTGGTCTGGGCGCTATGCAAGGAGTCTGGGATC 1003

Db 841 ATTGGTCCCCCTACAATTAGTACTGGCCACCAAGAAATACAGATGATCAGGACTC 900

Qy 1004 TGGGCTTACTACGAGATCTGTACCTTCTGAAATGAGGCGCTACGAGGATGGATGCC 1063

Db 901 CTGGCTTACTATGAGGTTTGTACATTTCTGAATGAAGGAGCGCACTGAGGTCTGGATGCC 960

Qy 1064 CCTCAGGAAGTGCCTTATGCTATCAGGCAATCTGTGGTGTGGCTATGACACATCAAG 1123

Db 961 CCCGGAAGTACCTATGCTATCAGGGAATAGTGGTGTGGTGTGGTATGACAAATGTCAGG 1020

Qy 1124 AGCTTCGATATTAAGGCTCAATGGCTTAAGCAACAATAATTTGAGGCGCATGGTCTGG 1183

Db 1021 AGCTTCAAGTTGAAGGCTCAGTGGCTCAAGGACAAATTTAGGAGGTGCGTGGTCTGG 1080

Qy 1184 GCAATGATCTGGATGACTTCACTGGCAGCTTTTCGCAACGAGGCGAAATTTCCGCTAATC 1243

Db 1081 CCCGTGACATGGATGACTTCACTGAGTGTCTTCTGTCCAGAGACATTTCCCTCTGACA 1140

Qy 1244 TCCACCTGAGGAGGCGCTCGGCTCAGAGTGCAGTTCGCAAGTTCGACGGCTCC 1293

Db 1141 TCTACTTAAGGGAGGATCTCAATATACAGTGCAGTTCGCAAGGCGCC 1190

RESULT 13

AY049765 1209 bp mRNA linear ROD 06-AUG-2002

LOCUS Mus musculus secreted protein precursor Ym2 mRNA, complete cds.

DEFINITION AY049765

ACCESSION AY049765.2 GI:22123906

VERSION

KEYWORDS

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE 1 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. Webb,D.C., McKenzie,A.N. and Foster,P.S. Expression of the Ym2 lectin-binding protein is dependent on interleukin (IL)-4 and IL-13 signal transduction: identification of a novel allergy-associated protein. J. Biol. Chem. 276 (45), 41969-41976 (2001).

JOURNAL 21551268

MEDLINE 11553626

PUBMED

REFERENCE 2 (bases 1 to 1209) Webb,D.C. and Foster,P.S. Direct Submission Submitted (02-AUG-2001) Biochemistry and Molecular Biology, John Curtin School of Medical Research, Australian National University, Off Mills Rd, Acton, Canberra, ACT 2601, Australia

AUTHORS Webb,D.C. and Foster,P.S. 3 (bases 1 to 1209) Direct Submission Submitted (06-AUG-2002) Biochemistry and Molecular Biology, John Curtin School of Medical Research, Australian National University, Off Mills Rd, Acton, Canberra, ACT 2601, Australia

TITLE Sequence update by submitter

JOURNAL On Aug 6, 2002 this sequence version replaced gi:15705156.

REMARK Location/Qualifiers

COMMENT 1..1209

FEATURES /organism="Mus musculus" /strain="BALB/c" /db_xref="taxon:10090" /note="isolated from the lung lavage from allergic mice" 1..1209 /note="expression regulated by IL-4 and IL-13 signaling via the IL-4R alpha subunit" /codon_start=1 /product="secreted protein precursor Ym2" /protein_id="AAL03953.2" /db_xref="GI:22123907"

CDS

/translation="MAKLIILVTGLAILNLVOLGSSYQLMCIYTSWAKDRPTGSGFKPG NIDPCLTHIYAFAGMKNIEIYLSQDLRDYEALNGLDNTELKTLAIGGWKFG PAPSSMVSTPQNRQTFIKSVIRFQYDFNEDGLNDWQYPCSRGSPKDKHLFSLVQ EMKAPFEESTLNIHPELLLTSCAGIDVTKSYKTPELSQSLDYTOVMTYDLHDPK NGTGENSPLYKSPYDICKSADLNVDSIIYWKDHGAASEKLIVGFYATGHTFILSDP SKNGIGDPTVSAGPYGPKYITNEQGLLAYFEICTELNEGATEIFDATQEPVYAILGNEW GYNVRSFKLKAQWLKDNLLGGAVVWPLMDDFSGSFCHQGRPLPTTTLLKRLDNVHSA SCKASVYRGEL"

sig_peptide 326 a 273 c 282 g 328 t

BASE COUNT 1..63

ORIGIN

Query Match 42.4%; Score 689.4; DB 10; Length 1209;

Best Local Similarity 73.6%; Pred. No. 2.5e-174;

Matches 879; Conservative 0; Mismatches 316; Indels 0; Gaps 0;

Qy 104 ATGACAAAGCTTATTCCTCCACAGGCTTGTCTCTTATATGATTTTGACAGCTCGGCTCT 163

Db 1 ATGGCCAAAGCTCATTTCTGTACAGAGTCTGCGCAATCTCTGAATGTACAGCTGGGATCT 60

Qy 164 GCCTACCAGCTGACATCTACTTCAACCACTGGGCCCACTACGCGCCAGTGGGCGGC 223

Db 61 TCCTACCAGCTGATGTGCTACTATACCACTGGGCTAAGACAGAGCCCAACAGAGGAGT 120

Qy 224 TTCATGCTGACAAATCGACCCCTGCTCTGTACCCACCTGATCTACGCTTTTGTCTGG 283

Db 121 TTCAAACTGTATAATTGACCCCTGCTGTACTCACCTGATCTATGCTTTTGTCTGG 180

Qy 284 AGCAGAACACAGAGATCACCACTCAATGAGACGATGTGACTCTCTACCAAGCTTC 343

Db 181 ATGAAGAATAATGAGATCACTTACTTAAGTGAGCAAGACTTGGTGTGACTATGAAGCATTA 240

Qy 344 AATGGCTGAAAATAAGAACAGCAGCTGAAAACCTCTCTGGCCATTTGGAGGCTGGAAC 403

Db 241 AATGGCTGAAACAGACAGGAACACTGAGCTAAAACCTCTCTGGCCATTTGGAGGATGGAAG 300

Qy 404 TTCGGGACTGCCCTTTTACTGCCATGGTTTCTACTCTCTGAGAAACCGCCAGACTTTCATC 463

Db 301 TTTGGACCTGCCCGTTTCACTGCTCTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360

Qy 464 ACCTCAGTCAATAAATTCCTGCGCCAGTATGAGTTTGTACGGCTGGAGCTTTGACTGGGAG 523

Db 361 AAGTCAGTTATCAGATTCCTTCGTCAATATACTTTGATGGCTCTCAACCTGGACTGGCAG 420

Qy 524 TACCTTGGCTCTCTGGGAGCCCTCTCTCAGSACAAGCATCTTCTACTGCTCTCTGGTGCAG 583

Db 421 TACCTTGGGCTCTCAGSAGGCCCTCTCTAAGSACAACATCTCTTCTAGTGTCTCTGGTGCAG 480

Qy 584 GAAATGGGTGAAGCTTTTGGAGGAGGCGCAACAGATCAACAAAGCCAGGCTGTAGGTC 643

Db 481 GAAATGGTGAAGCTTTTGGAGGAGGATCCACTTTGAACCAACATTTCAAGGCTGTACTC 540

Qy 644 ACTGTGCGCTAGCTGCTGGCATCTCCAATATCCAGTCTGGCTATGAGATCCCCCACTG 703

Db 541 ACTTCCACAGAGCTGGATTCATTGAGTAACTGAGTCTGGTACAGATCCCTGAACTG 600

Qy 704 TCACAGTACCTGGACTACATCCATGTCATGACCTAGCACCTCCATGCTCTCTGGGAGGCG 763

Db 601 TCTCAGTCTCTCGACTATATTCAGGTCATGACATATGATCTCCATGATCTCTAAGATGGC 660

Qy 764 TACATGGGAGAACAGCCCTCTTACAATACCCGACTGACACCGGCGAGCAACGCCCTAC 823

Db 661 TACATGGGAGAAATAGTCCCTCTTATAAATCTCCATATGACATTTGAAAGAGTGTGTAT 720

Qy 824 CTCATGTGATTTATCTGATGACTACTTGAAGGACAATGGAGCACCAGCTGAGAGCTC 883

Db 721 CTCATGTGATTTCAATTTATCTACTGTGAGGAGCACCAGGAGCAGCTTCTTGAGAGCTC 780

Qy 884 ATCGTTGGATTTCCCTACCTTGACACAACTTCTCATCTGAGCAACCCCTCCCAACTGGA 943

Db 781 ATTGTGGGATTTCCAGCATATGTCATACCTTTATCTCTGAGTGACCTCTCTAAGATGGA 840

Qy 944 ATTGGTGGCCCACTCTGCTGGTGTCTGCTGGGCGCTATGCCAAGGAGTCTGGGATC 1003

Db 841 ATAGGTGACCCCTACTGTTAGTCTGGACCACCGAGAAAGTACACAAATGAACAAGGACTC 900
QY 1004 TGGGCTTACTACGAGATCTGATCTTCTGAAATATGAGCCACCTCAGGATGGATGCC 1063
Db 901 CTGGCTTACTTTGAGATTGTTACATTTCTGAATGAAGGAGCCACTGAGATCTTTGATGCC 960
QY 1064 CCTCAGGAAGTCCCTTATGCCCTATCAGGCAATGTGTGGTTGGCTATGACACATCAAG 1123
Db 961 ACCCAGGAATACCTATGCTCTATCTGGTATGATGAGTGGTTGGTTATGACAAATGTCAGG 1020
QY 1124 AGCTTCGATPATTAAAGCTCAATGCTTAAGCAACAAATTTGAGGGCCCATGTGCTGG 1183
Db 1021 AGCTTCAAGTTCAAGGCTCAGTGGCTCAAGGACAAATTTAGGAGGTGCCGTGCTGG 1080
QY 1184 GCCATGATCTGGATGACTTCACTGGCCTTCTGCAACCCAGGGCAAGTTTCCCTTAATC 1243
Db 1081 CCCCTGGACATGGATGACTTCACTGCTGCTTCTGTCACCGGAGCTTTCCCTCTGACA 1140
QY 1244 TCCACCTGAAGAAGGCCCTCGGCTCGCAGAGTGAAGTTGCACGGCTCCAGCTC 1298
Db 1141 ACTACTTTAAAGAGATCTGATGTACACAGTGCAGTTGCAAGGCTCTTATC 1195

RESULT 14
LOCUS AY065557 1209 bp mRNA linear ROD 22-MAY-2002
DEFINITION Mus musculus putative secretory protein precursor (Ym2) mRNA,
complete cds.
ACCESSION AY065557
VERSION AY065557.1 GI:18086513
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Chang, N.-C.A.
Tissue-specific Expression of Ym2 Protein
Unpublished
2 (bases 1 to 1209)
Chang, N.-C.A.
Direct Submission
Submitted (05-DEC-2001) Institute of Microbiology & Immunology,
National Yang-Ming University, #155 Sec. 2 Li-Nong St., Beitou,
Taipei 112, Taiwan
Location/Qualifiers
1. .1209
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
/clone="pGEMT-ez-Ym2.s"
/tissue_type="stomach"
1. .1209
/gene="Ym2"
1. .1209
/codon_start=1
/product="putative secretory protein precursor"
/protein_id="AAL57751.1"
/db_xref="GI:18086514"
/translation="MAKLIIVTGLAILNVQLSSSYQLMCIYTSWAKDRPTGSEFKPG
NIDPCLTLLIYAFAGMKNNEITYLSEODLRDYENALGNKDRTELKTLAIGWKKFG
PAPFSSMSTPONROTFIKSVIRFLROYNFDGLNLDWQYPPGRGPPDKHLFSLVQ
ENKRAFEESTLNHLPRLLTSTGAGFDIVKSGKIPELSLSLDYIQWTDLDHDPK
NGYGENSPLYKSDYIGKSLADNLVDIITYWKDHGAASEKLIIVFPAYGHTFLLSDP
SKNGIGDTPVSAQPKYTNQEQGLLAYFEICFLNEGATEIFEDAEQVFPYVILGNEW
GYDNRVSFKLKAQWLKNNLGLAVVWPLDMDDFSGSFCHQGRFPLTTTLKRLNLVHSA
SCKASYRGEL"

BASE COUNT 326 a 274 c 282 g 327 t
ORIGIN
Query Match 42.3%; Score 687.8; DB 10; Length 1209;

Best Local Similarity 73.5%; Pred. No. 6.7e-174;
Matches 878; Conservative 0; Mismatches 317; Indels 0; Gaps 0;
QY 104 ATGCAACAAGCTTATTTCTCTCTCACAGGTCTTGTCTTATATCTATCTGAAATTTGCAAGCTCGCTCT 163
Db 1 ATGGCCCAAGCTCATTTCTTGTCTCACAGGTCTGGCAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 60
QY 164 GCTTACCAGCTGACATGCTTACCAACTGGGCCACAGTACCGGCCAGGCTTGGGGCCG 223
Db 61 TCTTACCAGCTGATGTCTTCTTATACCAGCTGGGCTAAGGACAGGCCAACACAGAGGAGT 120
QY 224 TTTATGCTGCTGACAACATCGACCCCTGCTCTGTACCCACCTGATCTACCCCTTTGCTGGG 283
Db 121 TTCAAACTGGTAAATATTTGACCCCTGCTGTCTACTCACTGATCTATGCTTCTTCTTCTTCTTCTTCT 180
QY 284 AGGCAAGACAGAGATCACCACTACCACTCGAATGGAAGAGTGTACTCTCTACCAAGCTTTC 343
Db 181 ATGAAGATTAATGATCTTCTTACTTAAGTGAAGAGACTTGGTGAAGCTATGAAGCATTA 240
QY 344 AATGGCTGAAAAATAAGAAACAGCAGCTGAAACCTCTCTGGCCATTTGAGGCTGGAAC 403
Db 241 AATGGTCTGAAGACAGGAACTGAGCTTAAACTCTCTGGCCATTTGGAGATGGAAG 300
QY 404 TTCGGACTGCCCTTTCTACTGCCATGTTTCTACTCTCTGAGAACCGCCAGACTTTTCATC 463
Db 301 TTTGGACCTGCCCGTTTCCAGTTCCTACTCTCTCAGAACCGTCTCAGACATTCAT 360
QY 464 ACTCAGTCACTAAATTTCTGCGCAGTATGATGATTTGACGGCTGAGCTTTGACTGGGAG 523
Db 361 AAGTCAGTTATCAGATTCTCTGCTCAATATAACTTTTGTATGGCTTCAACCTGGAGTGGCAG 420
QY 524 TACCCTGGCTCTGCTGGGAGCCCTCTCAGGACCAAGCATCTCTTCACTCTCTCTCTCTCTCTCTCTCT 583
Db 421 TACCCTGGCTCTGAGGAAGCCCTCTCTAAGGACAACTCTCTTCACTCTCTCTCTCTCTCTCTCTCTCT 480
QY 584 GAAATGCTGAAAGCTTTTGGAGCAGGAGCCCAAGAGATCAACAAGCCCGAGGTGTATGGTCT 643
Db 481 GAAATGCTGAAAGCTTTTGGAGGAAGAAATCCACTTTTGAACACACATTTCAAGAGGCTGCTACTC 540
QY 644 ACTGCTGAGTACTGCTGGCATCTCCAATATCCAGTCTGGCTATGAGATCCCCCACTG 703
Db 541 ACTTCCACAGGAGCTGGATTTCATGACGTAATCAAGTCTGGGTACAAGTCCCTGAACTG 600
QY 704 TCACAGTACTCTGAGTACTACATCCATGCTACGACTACGACTCCATCGCTCTGGGAGGCT 763
Db 601 TCTCAGTCTCTGCTATATTTTCAAGTCTGATGATGATCTCCATGATCTTCAAGATGGC 660
QY 764 TACACTGGAGAACAGCCCCCTCTTACAAATACCCGACTGACACCGGACAGCAAGCCTAC 823
Db 661 TACACTGGAGAAAAATAGTCCCTCTATAAATCTCCATATGACATTTGGAAGAGTGTCTGAT 720
QY 824 CTCAATGTGATTTGCTCATGACTGAAAGCAATGGAGGACCACTGAGGAGCTGAGAGCTC 883
Db 721 CTCAATGTGATTTCAATTTATACCTACTGGAAGGACCATGGGAGCTCTTCTGAGAAGCTC 780
QY 884 ATCGTGGATTCCTACCTATGACACAACTTCTATCTCTGAGCAACCCCTCCACACCTGGA 943
Db 781 ATTTGGGATTTCCAGCATATGCTTATCTCTGAGTCACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 840
QY 944 ATTTGGTCCCCACCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1003
Db 841 ATAGGTGACCCCTACTGTTAGTCTGGACCACCGAGAAAGTACACAAATGAACAAGGACTC 900
QY 1004 TGGGCTTACTACGAGATCTGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1063
Db 901 CTGGCTTACTTTGAGATTTGTATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 960
QY 1064 CTTCAAGAGTCTCTTATGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1123
Db 961 ACCCAGGAAGTACCCCTATGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
QY 1124 AGCTTCTGATATTAAGGCTCAATGGCTTAAAGCAACAAATTTGGAGGCCCTCTGCTGCTGCTGCTGCT 1183

Db 1021 AGCTTCAAGTTGAAGGCTCAGTGGCTCAAGGACAACAATTTAGGAGTGCGGTGGTCTGG 1080

Qy 1184 GCCATTGATCTGGATGACTTCACTGGCAGTCTTCTGCAACAGGCAAGTTTCCCTTAATC 1243

Db 1081 CCCTGGACATGGATGACTTCACTGGTCTTCTGTCACAGGAGCCTTCCCTCTGACA 1140

Qy 1244 TCACCCCTGAAGAGGCGCTCGGCTCGAGAGTCAAGTTGCAAGGCTCCAGCTC 1298

Db 1141 ACTACTTTAAAGAGAGATCTGAATGTACACAGTCAAGTTGCAAGGCTCTTATC 1195

RESULT 15

AR206042

LOCUS AR206042

DEFINITION Sequence 3 from patent US 6372212.

ACCESSION AR206042

VERSION AR206042.1 GI:21504526

KEYWORDS

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 1656)

AUTHORS Gray,P.W.

TITLE Chitnase materials and methods

JOURNAL Patent: US 6372212-A 3 16-APR-2002;

FEATURES

Location/Qualifiers

1..1656

/organism="unknown"

BASE COUNT 365 a 497 c 447 g 347 t

ORIGIN

Query Match 29.7%; Score 483; DB 6; Length 1656;

Best Local Similarity 60.2%; Pred. No. 9.7e-119;

Matches 870; Conservative 0; Mismatches 545; Indels 30; Gaps 3;

Qy 84 GACTGGTGTGCTGACTGCAACATGACAAAGCTTATTCCTCAGAGTCTTGTCTATAC 143

Db 7 GCCTGGCGTGTGAGTGCATATGCTGGTCTGTGGCTGGGAGGTTTTCATGGTCTGC 66

Qy 144 TGAATTGACGCTGGCTCTGCTTACCAGTGCATCTACTTACCAACTGGGCCAGT 203

Db 67 TGATGATCCCATGGGCTCTGTCGAAACTGTGTCTACTTCAACCACTGGGCCAGT 126

Qy 204 ACCGGCAGCGCTGGGCGCTTCATGCTGACAACTGACCGCTGCTCTGTACCCACC 263

Db 127 ACAGACAGGGGAGGCTCGCTTCTGCCCAAGGACTTGGACCCAGCCTTTGCACCCACC 186

Qy 264 TGATCTACGCTTGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 323

Db 187 TCATCTACGCTTGTGGGATGACCAACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 246

Qy 324 TGACTCTTACCAAGCTTTCAATGGCTGAAATAAGAACAGCAGCTGAAACACTCTCC 383

Db 247 AGACTCTTACCAAGGAGTTCAATGGCTGAAAGAGTGAATCCCAAGCTGAAGACCTGT 306

Qy 384 TGGCCATTTGAGGCTGGAACCTTCGGAGCTGCCCTTTTCACTGCCATGGTTTCTACTCTG 443

Db 307 TAGCCATCGGAGGCTGGAATTTTCCAGCCTCAGAGTTTCAAGATATGGTACCCAGGCCA 366

Qy 444 AGAACGCCAGACTTTTCATCACCCTCAGTATCAAAATTCCTGGCCAGTATGAGTTTGAAG 503

Db 367 ACAACGCTCAGACCTTTGTCAACTCGGCCATCAGGTTTCTGCGCAATACAGCTTTGAGC 426

Qy 504 GGCTGGACTTTGACTGGGAGTACCTTGGCTCTCGTGGGAGCCCTCTCCAGCAAGCATC 563

Db 427 GCCTTGACCTTGAAGGAGTACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 486

Qy 564 TCTTCTACTGCTGGTGCAGGAAATCGTGAAGCTTTTGGAGAGGAGGAGGAGGAGGAGGAGG 623

Db 487 GCTTCAACACCTGGTACAGGACTTGGCCAAATGCCTTCCAGAGGAGGAGGAGGAGGAGGAGG 546

Qy 624 ACAAGCCAGGCTGATGGTGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 683

Search completed: July 3, 2003, 07:21:38
Job time : 4281.04 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 2, 2003, 22:31:11 ; Search time 372.46 Seconds
(without alignments)
9825.200 Million cell updates/sec

Title: US-10-004-219B-2

Perfect score: 1625

Sequence: 1 gcttccagctgtgtgta.....ctgcaataaaatcagcagtc 1625

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_101002:*

- 1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
- 2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
- 3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
- 4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
- 5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
- 6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
- 7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
- 8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
- 9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
- 10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
- 11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
- 12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
- 13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
- 14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
- 15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
- 16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
- 17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
- 18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
- 19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
- 20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
- 21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
- 22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
- 23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
- 24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1529.6	94.1	1678	22	AAH42025
2	1366.4	84.1	1368	22	AAH42013
3	1266.2	77.9	1369	24	ABN59993
4	701.8	43.2	1469	22	AAH42023
5	629.2	38.7	746	22	AAK88635
6	628.4	38.7	638	22	ABK11713
7	483	29.7	1656	19	AAV10436
8	483	29.7	1656	20	AAZ21848
9	483	29.7	1656	22	AAO3760

10	483	29.7	1656	24	ABL57381	Human chitinase cd
11	482	29.7	1768	18	AAT89181	Human chitotriosid
12	481.4	29.6	1643	18	AAT50833	Human chitinase cd
13	481	29.6	1636	19	AAV10435	Human chitinase cl
14	481	29.6	1636	20	AZ21847	MO-218 clone of hu
15	481	29.6	1636	22	AAO3759	Human chitinase cd
16	481	29.6	1636	24	ABL57380	Human chitinase cd
17	477.2	29.4	1713	18	AAT50834	Human chitinase cd
18	460.6	28.3	1637	18	AAT89180	Human chitotriosid
19	369	22.7	1432	19	AAV21689	DNA encoding a hum
20	369	22.7	1433	19	AAV13925	Human cartilage gp
21	364.4	22.4	1496	18	AAT97127	Human cartilage g1
22	358	22.0	1594	19	AAV21688	DNA encoding a hum
23	354.8	21.8	1526	19	AAV13926	Human cartilage gp
24	354.8	21.8	1594	19	AAV21687	DNA encoding a hum
25	350.8	21.6	2504	16	AAO90443	Murine oviduct spe
26	348.4	21.4	1391	24	AAO62615	cDNA sequence #402
27	347.4	21.4	1994	16	AAO90442	Bovine oviduct spe
28	344.4	21.2	2366	16	AAO90444	Hamster oviduct sp
29	342.6	21.1	1152	18	AAT99452	Bovine whey protei
30	339.8	20.9	1681	16	AAO85245	YKL-40 gene. Homo
31	339.8	20.9	1681	21	AAZ94901	Human cancer marke
32	302.2	18.6	2346	23	ABL03597	Drosophila melanog
33	281.2	17.3	1474	22	AAH23078	Osteoarthritis tis
34	263	16.2	17368	23	ABL03596	Drosophila melanog
35	226	13.9	966	9	AAH81756	Gene encoding poly
36	205	12.6	13497	23	ABL03557	Drosophila melanog
37	162.8	10.0	2089	23	ABL29841	Drosophila melanog
38	156	9.6	2452	18	AAH23557	Manduca sexta larv
39	154.8	9.5	288	24	ABN76932	Human glycoprotein
40	154.2	9.5	503	21	AAO0233	Human secreted pro
41	147.6	9.1	1433	23	ABL15099	Drosophila melanog
42	133.6	8.2	379	21	AAO0135	Human secreted pro
43	130.8	8.0	369	22	AAH42024	Disease treatment
44	130	8.0	1302	23	ABL13745	Drosophila melanog
45	130	8.0	3302	23	ABL19744	Drosophila melanog

ALIGNMENTS

RESULT 1
AAH42025
ID AAH42025 standard; DNA; 1678 BP.
XX
AC AAH42025;
XX
DT 24-AUG-2001 (first entry)
DE Disease treatment related oligonucleotide SEQ ID NO: 16.
DE Disease treatment; infection; chronic occlusive pulmonary disease;
KW bronchial asthma; ds.
XX
OS Homo sapiens.
XX
PN WO200136633-A1.
XX
PD 25-MAY-2001.
XX
PF 14-NOV-2000; 2000WO-JP08015.
XX
PR 15-NOV-1999; 99JP-0324467.
XX
PA (TAKE) TAKEDA CHEM IND LTD.
XX
PI Nakanishi A, Morita S;
XX
DR WPI; 2001-397791/42.
XX
PT New proteins, peptides and DNA for treatment of bronchial asthma,
XX chronic occlusive lung disease and infectious disease -

PS Example 5; Page 107-108; 114pp; Japanese.

The present invention provides the sequence of a protein which can be used in the treatment and prevention of infectious diseases. Inhibitors of the protein can be used to treat bronchial asthma and chronic occlusive pulmonary disease. The present sequence is an oligonucleotide described in the exemplification of the invention.

SQ Sequence 1678 BP; 397 A; 497 C; 418 G; 366 T; 0 other;

Query Match	94.1%	Score 1529.6;	DB 22;	Length 1678;
Best Local Similarity	99.7%;	Pred. No. 0;		
Matches 1532;	Conservative	0;	Mismatches	4;
			Indels	0;
			Gaps	0;

QY	90	TGCTGACTGCAACCATGACAAAGCTTATCTCTCTACAGGTCCTTATCTATACTGAATT	149
Db	129	TGCTCCCTGCACCCATGACAAAGCTTATCTCTCTACAGGTCCTTATCTATACTGAATT	188
QY	150	TGCAGCTCGGCTCTGGCTACACGTGACATGCTACTTACCAACTGGGCCGACGACCGC	209
Db	189	TGCAGCTCGGCTCTGGCTACACGTGACATGCTACTTACCAACTGGGCCGACGACCGC	248
QY	210	CAGGCTTGGGGCGCTTCATGTGCTGACAACATGCAACCCCTGCTCTGTACCCACCTGATCT	269
Db	249	CAGGCTTGGGGCGCTTCATGTGCTGACAACATGCAACCCCTGCTCTGTGTACCCACCTGATCT	308
QY	270	AGCCCTTTTCTGCTGGAGGCGAGAACACGAGATCACCACCATCGAATGGAACGATGTGACTC	329
Db	309	AGCCCTTTTCTGCTGGAGGCGAGAACACGAGATCACCACCATCGAATGGAATGTGTGACTC	368
QY	330	TCTACCAAGCTTTCAATGGCCTGAAAAATAAGAACACGACGCTGMAAACHCTCTCTGGCCA	389
Db	369	TCTACCAAGCTTTCAATGGCCTGAAAAATAAGAACACGACGCTGMAAACHCTCTCTGGCCA	428
QY	390	TTGGAGGCTTGGAACTTTGGGAGCTGCCCTTTTCACTGCCATGGTTTCTACTCTCTGAGAACC	449
Db	429	TTGGAGGCTTGGAACTTTGGGAGCTGCCCTTTTCACTGCCATGGTTTCTACTCTCTGAGAACC	488
QY	450	GCAGAGCTTTCAATCACTCAGTCATCAAAATTCCTGCGCCAGTATGAGTTTTCAGGGGCTGG	509
Db	489	GCAGAGCTTTCAATCACTCAGTCATCAAAATTCCTGCGCCAGTATGAGTTTTCAGGGGCTGG	548
QY	510	ACTTTGACTGGGAGTACCTGGCTCTCTGCTGGGAGCCCTCTCTAGGACAAGCATCTCTTCA	569
Db	549	ACTTTGACTGGGAGTACCTGGCTCTCTGCTGGGAGCCCTCTCTAGGACAAGCATCTCTTCA	608
QY	570	CTGTCTGTGTGAGGAAATGCGTGAAGCTTTTTCAGCAGGAGGCCAAGCAGATCAACAAGC	629
Db	609	CTGTCTGTGTGAGGAAATGCGTGAAGCTTTTTCAGCAGGAGGCCAAGCAGATCAACAAGC	668
QY	630	CCAGGCTGATGGTCACTGCTGCAGTAGCTGTCTGGCATCTCCAAATATCCAGTCTGGCTATG	689
Db	669	CCAGGCTGATGGTCACTGCTGCAGTAGCTGTCTGGCATCTCCAAATATCCAGTCTGGCTATG	728
QY	690	AGATCCCCCAACTGTCCACAGTACCTGGACTACATCCATGTCATGACCTAGGACCTCCATG	749
Db	729	AGATCCCCCAACTGTCCACAGTACCTGGACTACATCCATGTCATGACCTAGGACCTCCATG	788
QY	750	GCTCTCTGGAGGGCTACACTGGAGAGAACAGCCCCCTCTACAAATATCCGACCTGACACCG	809
Db	789	GCTCTCTGGAGGGCTACACTGGAGAGAACAGCCCCCTCTACAAATATCCGACCTGACACCG	848
QY	810	GCAGCAACGCCCTACCTCAATGTGGATTATGTCAATGAACCTACTGGGAAGGACAATGGAGCAC	869
Db	849	GCAGCAACGCCCTACCTCAATGTGGATTATGTCAATGAACCTACTGGGAAGGACAATGGAGCAC	908
QY	870	CAGCTTGAGAAGCTCATCGTGTGGATTCCCTACCTATGTGCACACAACTTCATCCTGAGCAACC	929
Db	909	CAGCTTGAGAAGCTCATCGTGTGGATTCCCTACCTATGTGCACACAACTTCATCCTGAGCAACC	968
QY	930	CCTCCACACATGGAATGTGTGGCCCCCACTCTGGTGCTGGTCTCTGTCTGGGGCCCTATGCCA	989
Db	969	CCTCCACACATGGAATGTGTGGCCCCCACTCTGGTGCTGGTCTCTGTCTGGGGCCCTATGCCA	1028

QY	990	AGGAGTC	GGGATCTGGCCTTACTACGAGATCTGTACTCTTCTGAAAAATGGAGCACTC	1049
Db	1029	AGGAGTC	GGGATCTGGCCTTACTACGAGATCTGTACTCTTCTGAAAAATGGAGCACTC	1088
QY	1050	AGGATGGGATGCCCTCAGGAAGTGCCCTTATGCCATATCAGGGCAATGTGTGGTTGGCT	1109	
Db	1089	AGGATGGGATGCCCTCAGGAAGTGCCCTTATGCCATATCAGGGCAATGTGTGGTTGGCT	1148	
QY	1110	ATGACAACATCAAGAGCTTCGATATTAAAGCTCAATGGCTTAAAGCACACAATAATTTGGAG	1169	
Db	1149	ATGACAACATCAAGAGCTTCGATATTAAAGCTCAATGGCTTAAAGCACACAATAATTTGGAG	1208	
QY	1170	CGCCATAGTCTGGGCCATTGATCTGATGACTTCACTTGGACATTTCTGCAACCAAGGGCA	1229	
Db	1209	CGCCATAGTCTGGGCCATTGATCTGATGACTTCACTTGGACATTTCTGCAACCAAGGGCA	1268	
QY	1230	AGTTTCCCCTTAATCTCCACCCTGAAGAGGCCCTCGCCCTGCAGATGCAAGTTGCACGG	1289	
Db	1269	AGTTTCCCCTTAATCTCCACCCTGAAGAGGCCCTCGCCCTGCAGATGCAAGTTGCACGG	1328	
QY	1290	CTCCAGCTCAGCCCCATTGAGCCAAATACTGCTGCTCCAGTGGCAGCGGGAACGGGAGCG	1349	
Db	1329	CTCCAGCTCAGCCCCATTGAGCCAAATACTGCTGCTCCAGTGGCAGCGGGAACGGGAGCG	1388	
QY	1350	GGAGTAGAGCTCTGGAGGAGCTCGGGAGGCACTGGATTCGTGTGTCTGAGAGCCAACG	1409	
Db	1389	GGAGTAGAGCTCTGGAGGAGCTCGGGAGGCACTGGATTCGTGTGTCTGAGAGCCAACG	1448	
QY	1410	GCCTTACCCCGTGGCAATAACAGAAATGCCCTTCTGGCACTCGCTGAATGGAGTCACGT	1469	
Db	1449	GCCTTACCCCGTGGCAATAACAGAAATGCCCTTCTGGCACTCGCTGAATGGAGTCACGT	1508	
QY	1470	ACCAAGCAGAACTGCAGCGCGGCTTCTCTTCACACCAGCTGTGATTGCTGCAACTGGG	1529	
Db	1509	ACCAAGCAGAACTGCAGCGCGGCTTCTCTTCACACCAGCTGTGATTGCTGCAACTGGG	1568	
QY	1530	CATAAAGCTGACCTGGTCTATATCCCTAGAGTTCAGTCTCTTTTGTAGACATGTT	1589	
Db	1569	CATAAAGCTGACCTGGTCTATATCCCTAGAGTTCAGTCTCTTTTGTAGACATGTT	1628	
QY	1590	GCCCTACCTAAAGTCTCTGCAATAAAATCAGCAGTC	1625	
Db	1629	GCCCTACCTAAAGTCTCTGCAATAAAATCAGCAGTC	1664	
RESULT 2				
AAH42013				
ID	AAH42013 standard; DNA; 1368 BP.			
AC	AAH42013;			
DT	24-AUG-2001 (first entry)			
DE	Disease treatment related oligonucleotide SEQ ID NO: 3.			
KW	Disease treatment; infection; chronic occlusive pulmonary disease;			
KW	bronchial asthma; ds.			
OS	Homo sapiens.			
PN	W0200136633-A1.			
PD	25-MAY-2001.			
PF	14-NOV-2000; 2000WO-JP08015.			
PR	15-NOV-1999; 99JP-0324467.			
PA	(TAKE) TAKEDA CHEM IND LTD.			
PI	Nakanishi A, Morita S;			

DR	WPI; 2001-397791/42.	
XX	New proteins, peptides and DNA for treatment of bronchial asthma,	
PT	chronic occlusive lung disease and infectious disease -	
XX	Claim 5; Page 100; 114pp; Japanese.	
PS		
XX	The present invention provides the sequence of a protein which can be	
CC	used in the treatment and prevention of infectious diseases. Inhibitors	
CC	of the protein can be used to treat bronchial asthma and chronic	
CC	occlusive pulmonary disease. The present sequence is an oligonucleotide	
CC	described in the exemplification of the invention.	
XX		
SQ	Sequence 1368 BP; 324 A; 398 C; 349 G; 297 T; 0 other;	
	Query Match 84.1%; Score 1366.4; DB 22; Length 1368;	
	Best Local Similarity 99.9%; Pred. No. 0;	
	Matches 1367; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
Qy	167 TACAGCTGACATGCTACTTACCAACTGGGCCAGTACCGGCCAGCCCTGGGCGGCTTC 226	
Db	1 TACCAGCTGACATGCTACTTACCACCTGGGCCAGTACCGGCCAGCCCTGGGCGGCTTC 60	
Qy	227 ATGCTTACAAACATCGACCCCTGCCTCTGTACCCACCTGATCTACGCCCTTTCCTGGGAGG 286	
Db	61 ATGCTTACAAACATCGACCCCTGCCTCTGTACCCACCTGATCTACGCCCTTTCCTGGGAGG 120	
Qy	287 CAGAAACAGAGATCACACCATCGAATGGAACGATGTGACTCTCTACCAAGCTTTCAT 346	
Db	121 CAGAAACAGAGATCACACCATCGAATGGAATGTGACTCTCTACCAAGCTTTCAT 180	
Qy	347 GGCCTGAAAATAGAACAGCCAGCTGAAACTCTCTGGCCATTGGAGGCTTGGAACTTC 406	
Db	181.GGCCTGAAAATAGAACAGCCAGCTGAAACTCTCTGGCCATTGGAGGCTTGGAACTTC 240	
Qy	407 GGGACTGCCCTTTCACCTGCCATGGTTTCTACTCTCCTGAGAACCCGAGACTTTCATCACC 466	
Db	241 GGGACTGCCCTTTCACCTGCCATGGTTTCTACTCTCCTGAGAACCCGAGACTTTCATCACC 300	
Qy	467 TCAGTCAATCAAAATCCCTGGGCCAGTATGAGTTTGACGGCTGGACTTTGACTGGGAGTAC 526	
Db	301 TCAGTCAATCAAAATCCCTGGGCCAGTATGAGTTTGACGGCTGGACTTTGACTGGGAGTAC 360	
Qy	527 CCTGGCTCTCGTGGAGCCCTCCCTCAGGACAAGCATCTCTTCACTGCTCTGGTCAGAGAA 586	
Db	361 CCTGGCTCTCGTGGAGCCCTCCCTCAGGACAAGCATCTCTTCACTGCTCTGGTCAGAGAA 420	
Qy	587 ATGGCTGAAGCTTTTGAGCAGGAGGCCAAGCAGATCAACAAAGCCAGCTGATGGTCACT 646	
Db	421 ATGGCTGAAGCTTTTGAGCAGGAGGCCAAGCAGATCAACAAAGCCAGCTGATGGTCACT 480	
Qy	647 GCTGAGTAGCTGCTGGCATCTCCAATATCCAGTCTGGCTATGAGATCCCCCAACTGTCA 706	
Db	481 GCTGAGTAGCTGCTGGCATCTCCAATATCCAGTCTGGCTATGAGATCCCCCAACTGTCA 540	
Qy	707 CAGTACCTGGACTACATCCATGTGATGACCTACGACCTCCATGGCTCCTGGGAGGGCTAC 766	
Db	541 CAGTACCTGGACTACATCCATGTGATGACCTACGACCTCCATGGCTCCTGGGAGGGCTAC 600	
Qy	767 ACTGGAGAGAACAGCCCCCTCTTCAAAATACCGACTGACACCGGCACGAACGCTTACCTC 826	
Db	601 ACTGGAGAGAACAGCCCCCTCTTCAAAATACCGACTGACACCGGCACGAACGCTTACCTC 660	
Qy	827 AATGTGGATTATGTGATGAATCACTGGAAGGACAATGGACACCAGCTGAGAGCTCATC 886	
Db	661 AATGTGGATTATGTGATGAATCACTGGAAGGACAATGGACACCAGCTGAGAGCTCATC 720	
Qy	887 GTTGGATTCCCTACCTATGGACACAACCTTCATCTGTAGCAACCCCTCCAACTCTGGAATT 946	
Db	721 GTTGGATTCCCTACCTATGGACACAACCTTCATCTGTAGCAACCCCTCCAACTCTGGAATT 780	
Qy	947 GGTGCCCCCCACCTCTGGTGTGCTGTCTGGCCCTATGCCAAGGAGTCTGGGATCTGG 1006	

CC used in the treatment and prevention of infectious diseases. Inhibitors
CC of the protein can be used to treat bronchial asthma and chronic
CC occlusive pulmonary disease. The present sequence is an oligonucleotide
CC described in the exemplification of the invention.

XX
SQ Sequence 1469 BP; 406 A; 331 C; 332 G; 400 T; 0 other;

Query Match 43.2%; Score 701.8; DB 22; Length 1469;

Best Local Similarity 74.3%; Pred. No. 1.3e-188;

Matches 886; Conservative 0; Mismatches 307; Indels 0; Gaps 0;

QY	101	ACCATGACAAAGCTTATTCCTCCTACAGCTCTGTGCTTATCTACTGAATTTGACGCTCGGC	160
Db	6	ACCATGGCAAGCTCAATTTCTGTCACAGCTCTGGCAATTTCTTGAACGCTACAGCTGGGA	65
QY	161	TCTGCTACAGCTGACATGCTACTTACCACTGGCCGACGTAACGCGGCGAGCCCTGGG	220
Db	66	TCTTCTACAGCTGATGTCTACTATACAGTTGGCTAAGGACAGGCAATAGAGGG	125
QY	221	CGTTCATGCTGACAAATGAGCAAGCCGCTGCTGTACCCACCTGATCAGCCCTTGTCT	280
Db	126	AGTTTCAAACTGGTAATATGACCCCTGCTGTGCTACCTGATCTATGCTTGTCT	185
QY	281	GGGAGGCAACACAGATACCCACCATCGAATGGAACGATGTGACTCTCTACCAAGCT	340
Db	186	GGAATCGAATAATGAGATCACTTACACACATGAGCAAGACTTGGCTGACTATGAGCA	245
QY	341	TTCAATGGCTGAAAAATAGAACAGCCAGCTGAAACCTCTCTGCGCCATTTGAGGCTGG	400
Db	246	TGGAATGGTCTGNAAGACAGACACTGAGCTTAAACACTCTCTGCGCCATTTGAGGATGG	305
QY	401	AACTTCGGAGCTGCCCTTTCACTGCTGCTTCTACTCTCTGAGAACGCCAGACTTTC	460
Db	306	AACTTTGGACCTGCCCGCTTCACTGCTGCTTCTACTCTCTCAGAACCTCAGATATTC	365
QY	461	ATCACTCAGTCATCAATCTCTGCGCAGTATGAGTTGACGGCTGGACTTTGACTGG	520
Db	366	ATTCACTCAGTTATCAATCTCTGCTGCTCAATATCACTTTGATGGCTCAACCTGGACTGG	425
QY	521	GAGTACCTTGGCTCTGCTGGAGCCCTCTCAGGACAAAGCTCTCTCACTGCTCCCTGGT	580
Db	426	CAGTACCTTGGCTCTGAGAGAGCCCTCTCAGGACAAACATCTCTTCACTGCTTGGTG	485
QY	581	CAGGAATCGTGAAGCTTTTGAGCAGGAGGCAAGCAGATCAACAAAGCCGAGCTGATG	640
Db	486	AAGGAATCGTAAAGCTTTTGAGGAAGATCTGTGAGAACACATTTCAAGGCTGCTA	545
QY	641	GTCATGCTGAGTAGCTGCTGGCATCTCCAATATCCAGTCTGGCTATGAGATCCCCAA	700
Db	546	CTCACTTCCACAGGAGCAGGAATCAATTGACGTAATCAAGTCTGGGTACAAAGATCCCTGAA	605
QY	701	CTGTACACTACTGAGCTACATCTCCATGTCATGACCTAGCCTCCATGCTCTGGAG	760
Db	606	CTGTCTCAGCTCTGCTGATATTTAGGTCATGACATATGATCTCCATGATCTCAAGAT	665
QY	761	GGCTACACTGGAGAACAGCCCTCTCAATATCCCGACTGACACCGGCGAGCAACGCC	820
Db	666	GGCTACACTGGAGAAAATAGTCCCTCTATAATCTCCATATGACATTTGAAAGATGCT	725
QY	821	TACCTCAATGGATTTATGTCATGAACTACTGGAAGCAATGGAGCACCTGAGAG	880
Db	726	GATCTCAATGGATTTCAATCACTTCTCTGAGAGGACCATTTGAGCAGCTTCTGAGAAG	785
QY	881	CTCATCTGTTGGATTTCCCTACCTATGAGCAACATCTCATCTGAGCAACCTCCACACT	940
Db	786	CTCATTTGGATTTCCACATATGGGCATACCTTTATCTGATGACCTCTTCAAGAT	845
QY	941	GGAATTTGGCCCCACCTCTTGGTGTCTGCTGGGCCCTTATCCCAAGGAGTCTGGG	1000
Db	846	GGAATTTGGTCCCTTACATTTAGTACTGCCCCCAGGAAAGTACACAGATGAATCAGGA	905
QY	1001	ATCTGGCTTACTACGATCTGTACCTTCTCTGAAATATGGAGCCACTCAGGATGGGAT	1060

Db	906	CTCTCGCTTACTATGAGGTTTGTACATTTCTGAATGAAGAGCCACTGAGGCTCGGAT	965
QY	1061	GCCCTCTCAGAACTGCTTATGCTATCAGGGCAATGTGTGGTTGGCTATGACAATC	1120
Db	966	GCCCCCAGGAAGTACCCTATGCTATCAGGGTAATGAGTGGTTGGTTATGACATGTC	1025
QY	1121	AAGAGCTTCGATATTAAAGGCTCAATGGCTTAAGCACAACAATAATTTGAGGCGCATGGTC	1180
Db	1026	AGGAGCTTCAAGTTGAAGGCTCAGTGGCTCAAGGACAACAATTTAGGAGTGCCTGGTC	1085
QY	1181	TGGGCCATGATCTGATGACATTCACCTGGCAGCTTTCTGCAACAGGCAAGTTTCCCTTA	1240
Db	1086	TGCCCCCTGACATGATGACATTCAGTGGTTCTTCTGTCAACAGAGACATTTCCCTCTG	1145
QY	1241	ATCTCCACCTGAAGAAGGCCCTCGCCTGCAAGTGCAGTTGTCACGGCTCC	1293
Db	1146	ACATCTACTTTAAGGAGATCTCAATATACACAGTGCAGATTGCAAGGGCCC	1198

RESULT 5
AAK8635
ID AAK8635 standard; cDNA; 746 BP.
XX
AC AAK8635;
XX
DT 05-NOV-2001 (first entry)
XX Human digestive system antigen coding sequence SEQ ID NO: 951.
DE Human; digestive system antigen; gene therapy; cancer; appendicitis;
KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
KW digestive system disorder; Meckel's diverticulum; ss.
XX Homo sapiens.
XX
PN WO20015314-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01324.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216847.
PR 11-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 18-AUG-2000; 2000US-0225759.
PR 22-AUG-2000; 2000US-0226681.

950	QY	GC	CCCCACCTCTGGTGCTGCTGCTCTGCTGGGCCCTATGCCAAGGAGTCTGGGATCTGGGCT	1009
31	Db	GC	CTGCAGGTACCGGTCCGGAATTCCCGGGTCGACCCACGCGTCCGCTGGGATCTGGCT	90
1010	QY	TACT	ACGAGATCTGTACCTTCTCTGAAAAATGGAGCCACTCAGGGATGGGATGCCCTCAG	1069
91	Db	TACT	ACGAGATCTGTACCTTCTCTGAAAAATGGAGCCACTCAGGGATGGGATGCCCTCAG	150
1070	QY	GAAGTCCCTTATGCCCTATCAGGGCAATGTGTGGTTGGCTATGACAACATCAAGAGCTTC	1129	
151	Db	GAAGTCCCTTATGCCCTATCAGGGCAATGTGTGGTTGGCTATGACAACATCAAGAGCTTC	210	
1130	QY	GATATTAAAGCTCAATGGCTTAAGCACAACAAATTTGGAGCGCCATGGTCTGGGCCATT	1189	

Db 211 GATATTAAAGCTCAATGGCTTAAGCACAAATAATTGGAGGGCCAGCTGTCTGGGCCATT 270

QY 1190 GATCTGATGACTTCACTGGCACTTTCTGCAACAGGGCAAGTTTCCCTTAATCTCCACC 1249

Db 271 GATCTGATGACTTCACTGGCACTTTCTGCAACAGGGCAAGTTTCCCTTAATCTCCACC 330

QY 1250 CTGAAGAAGGCCCTCGGCCCTGCAGAGTGCAAGTTGACACGGCTCCAGCTCAGCCCAATTGAG 1309

Db 331 CTGAAGAAGGCCCTCGGCCCTGCAGAGTGCAAGTTGACACGGCTCCAGCTCAGCCCAATTGAG 390

QY 1310 CCAATAACTGCTGCTCCAGTGGCAGCGGAGAGCGGAGTAGAGCTCTGAGAGGC 1369

Db 391 CCAATAACTGCTGCTCCAGTGGCAGCGGAGAGCGGAGTAGAGCTCTGAGAGGC 450

QY 1370 AGCTCGGAGGAGTGATTTCTGCTGTGTCAGAGCAACGGCTCTACCCCGTGGCAAT 1429

Db 451 AGCTCGGAGGAGTGATTTCTGCTGTGTCAGAGCAACGGCTCTACCCCGTGGCAAT 510

QY 1430 AACAGAAATGCCCTTCTGGCACTGCGTGAATGGAGTCACTACACAGAGAACTGCCAGGCC 1489

Db 511 AACAGAAATGCCCTTCTGGCACTGCGTGAATGGAGTCACTACACAGAGAACTGCCAGGCC 570

QY 1490 GGGCTTGTCTGACACAGCTGTGATTCGCACTGGCGGATATAAAGCTGACCTGTCTA 1549

Db 571 GGGCTTGTCTGACACAGCTGTGATTCGCACTGGCGGATATAAAGCTGACCTGTCTA 630

QY 1550 TATTCCTAGAGTTCAGCTCTTTTGTCTAGACATGTGCGCTTACCTAAAGTCTCTG 1609

Db 631 TATTCCTAGAGTTCAGCTCTTTTGTCTAGACATGTGCGCTTACCTAAAGTCTCTG 690

QY 1610 AATAAATCAGCAGTC 1625

Db 691 AATAAATCAGCAGTC 706

RESULT 6

ID ABK11713

XX ABK11713 standard; cDNA; 638 BP.

AC ABK11713;

DT 05-JUN-2002 (first entry)

XX DNA encoding novel human secreted protein #5.

Secreted protein; gene therapy; immunostimulant; human; stroke;
Immune system disorder; immunologic deficiency syndrome; anaemia;
ataxia telangiectasia; HIV; human immunodeficiency virus; heart attack;
Wiskott-Aldrich disorder; thrombocytopenia; scarring; erythematosis;
autoimmune disease; Addison's disease; encephalomyelitis; lupus;
multiple sclerosis; autoimmune thyroiditis; diabetes; inflammation;
nephritis; ischaemia; hypergammaglobulinaemia; Crohn's disease;
sarcoidosis; Gaucher's disease; cardiovascular disease; telangiectasia;
persistent truncus arteriosus; arrhythmias; angina pectoris; tumour;
myocardial infarction; hypertension; neovascularisation; angiogenesis;
cancer; ocular angiogenic disease; hyperproliferative disorder;
diabetic retinopathy; uveitis; rheumatoid arthritis; psoriasis;
wound healing; endometriosis; vasculogenesis; atherosclerosis;
nervous system disease; Parkinson's disease; Alzheimer's disease;
infectious disease; gene; ss.

XX Homo sapiens.

Key	Location/Qualifiers
CD5	181..540
	/*tag= a
FT	/product= "Novel secreted protein"
FT	181..183
FT	/*tag= b
FT	184..537
FT	/*tag= c
FT	/label= mature_human_secreted_protein
XX	

PN WO200214341-A1.

XX 21-FEB-2002.

XX 13-AUG-2001; 2001WO-US25288.

XX 14-AUG-2000; 2000US-225215P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Barash SC, NI J, Ruben SM, Rosen CA, Shi Y;

PI WPI; 2002-257592/30.

DR P-PSDB; AAU77733.

XX Novel human secreted proteins useful for treating immune system e.g. anaemia, cardiovascular e.g. myocardial infarction, neurological disorders e.g. Alzheimer's disease, Parkinson's disease -

PT Claim 1; Page 322; 337pp; English.

XX The invention describes eight novel human secreted polypeptides (I). (I) and the polynucleotides encoding them are useful for treating, disorders, including disorders of immune system e.g. immunologic deficiency syndrome, ataxia telangiectasia, HIV infection (human immunodeficiency virus), Wiskott-Aldrich disorder, anaemia, thrombocytopenia, heart attacks (infarction), strokes, or scarring; autoimmune diseases e.g. Addison's disease, encephalomyelitis, multiple sclerosis, autoimmune thyroiditis, lupus, erythematosis, insulin dependent diabetes mellitus; disorders of inflammation e.g. nephritis, ischaemia-reperfusion injury; hypergammaglobulinaemia, Crohn's disease, sarcoidosis, Gaucher's disease; cardiovascular disease e.g. cardiovascular abnormalities, persistent truncus arteriosus, arrhythmias, angina pectoris, myocardial infarction, hypertension, telangiectasia, ischaemia includes cerebral ischaemia, late mucocutaneous lymph node syndrome; disorders and/or conditions associated with neovascularisation; cancers which involve angiogenesis, benign tumours, ocular angiogenic diseases, hyperproliferative disorders, diabetic retinopathy, uveitis, rheumatoid arthritis, psoriasis, delayed wound healing, endometriosis, vasculogenesis, atherosclerosis; nervous system diseases, degeneration associated with Parkinson's disease, a Alzheimer's disease and infectious diseases. This sequence encodes a novel human secreted protein, described in the invention.

XX 'Sequence 638 BP; 153 A; 171 C; 169 G; 145 T; 0 other;

Query Match 38.7%; Score 628.4; DB 24; Length 638;
Best Local Similarity 99.8%; Pred. NO. 5.9e-168;
Matches 629; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 996 CTGGGATCTGGGCTTACTACGAGATCTGTACCTTCTGAAAATGGAGCCTCAGGGAT 1055

Db 2 CTGGGATCTGGGCTTACTACGAGATCTGTACCTTCTGAAAATGGAGCCTCAGGGAT 61

QY 1056 GGGATCCCCCTCAGGAAGTGCCTTATGCCTATCAGGGCAATGTGGTGGCTATGACA 1115

Db 62 GGGATCCCCCTCAGGAAGTGCCTTATGCCTATCAGGGCAATGTGGTGGCTATGACA 121

QY 1116 ACATCAAGAGCTTCGATATTAAGGCTCAATGGCTTAAGCACACAATTTGGAGGCGCCA 1175

Db 122 ACATCAAGAGCTTCGATATTAAGGCTCAATGGCTTAAGCACACAATTTGGAGGCGCCA 181

QY 1176 TGGTCTGGGCCATTGATCTGGGACTTCTGACCTGGCCTTCTGCAACACCGGCAAGTTTC 1235

Db 182 TGGTCTGGGCCATTGATCTGGGACTTCTGACCTGGCCTTCTGCAACACCGGCAAGTTTC 241

QY 1236 CCCTAATCTCCACCCCTGAAGAGCCCTCGGCCCTGCAGAGTGCAGAGTGCACGGCTCCAG 1295

Db 242 CCCTAATCTCCACCCCTGAAGAGCCCTCGGCCCTGCAGAGTGCAGAGTGCACGGCTCCAG 301

QY 1296 CTCAGCCCCATTGAGCCCAATAACTGTGCTGCCAGTGGCAGCGGGAACGGAGCGGAGTA 1355

Db 302 CTCAGCCCCATTGAGCCCAATAACTGTGCTGCCAGTGGCAGCGGGAACGGAGCGGAGTA 361

QY 1356 GCAGCTGTGGAGGAGCTGGGAGGACAGTGTGATTTGTGTCGTGAGAGCCAAACGGCCCTCT 1415
 Db 362 GCAGCTGTGGAGGAGCTGGGAGGACAGTGTGATTTGTGTCGTGAGAGCCAAACGGCCCTCT 421
 QY 1416 ACCCGTGGGAAATACAGAAATGCTTCTGGGAGTGGTGAATGGAGTACGATACAGC 1475
 Db 422 ACCCGTGGGAAATACAGAAATGCTTCTGGGAGTGGTGAATGGAGTACGATACAGC 481
 QY 1476 AGAAGTGGGAGGAGTGTGTCGTGAGAGGAGTGTGATTTGTGTCGTGAGAGCCAAACGGCCCTCT 1535
 Db 482 AGAAGTGGGAGGAGTGTGTCGTGAGAGGAGTGTGATTTGTGTCGTGAGAGCCAAACGGCCCTCT 541
 QY 1536 CCTGAGCTGTGATTTATTCCTAGAGTTCAGTCTCTTTTCTGAGGACATGTTGCCCT 1595
 Db 542 CCTGAGCTGTGATTTATTCCTAGAGTTCAGTCTCTTTTCTGAGGACATGTTGCCCT 601
 QY 1596 ACCTAAAGTCTGCAATAAATACAGAGTCT 1625
 Db 602 ACCTAAAGTCTGCAATAAATACAGATTC 631

RESULT 7

AAV10436

ID AAV10436 standard; cDNA; 1656 BP.

XX AC

XX AAV10436;

XX 15-JUN-1998 (first entry)

XX Human chitinase clone MO-13B cDNA.

XX Chitinase; human; fungal infection; immunogen; diagnosis; treatment;
 KW Gaucher's disease; transgenic; detection; hybridisation; antifungal;
 KW rheumatoid arthritis; overexpression; extracellular matrix; ss.

XX Homo sapiens.

XX Key Location/Qualifiers
 FT CDS 27..1427

FT sig_peptide /*tag= a

FT /*tag= b

FT mat_peptide /*tag= c

FT /*product= chitinase

FT /*note= "from clone MO-13B"

XX W09747752-A1.

XX 18-DEC-1997.

XX 16-JUN-1997; 97WO-US10460.

XX 14-JUN-1996; 96US-0663618.

XX (ICOS-) ICOS CORP.

XX Gray PW;

XX WPI: 1998-052316/05.

XX P-PSDB; AAW40260.

XX Nucleic acids encoding human chitinase - useful as antifungal

XX agents, especially in combination with other antifungals

XX Claim 9; Page 42-44; 63pp; English.

XX This sequence encodes a novel human chitinase isolated from clone MO-13B.

XX Chitinases are useful for treating or preventing fungal infection and

XX as immunogens for generating antibodies which are used to purify, detect

XX and quantify chitinases, e.g. for diagnosis of Gaucher's disease. The

XX nucleic acid sequence of the chitinase is also useful as a probe to

XX identify and isolate genomic DNA encoding chitinases or similar proteins.

CC or cells expressing them or to generate transgenic ('knockout') rodents.
 CC It can also be used in hybridisation assays and to detect genetic
 CC alterations in the chitinase gene related to disease. Agents that inhibit
 CC this protein may be useful in treatment of Gaucher's disease and
 CC rheumatoid arthritis, where overexpression of the protein can damage
 CC the extracellular matrix. Chitinase also improves the activity of other
 CC antifungal agents and may allow a reduction in the dose of such agents,
 CC and thus of their side effects.

XX Sequence 1656 BP; 365 A; 497 C; 447 G; 347 T; 0 other;

Query Match 29.7%; Score 483; DB 19; Length 1656;
 Best Local Similarity 60.2%; Pred. No. 1.8e-126;
 Matches 870; Conservative 0; Mismatches 545; Indels 30; Gaps 3;

QY 84 GACTGGTGTGCTGACGTCACCAACATGACAAAGCTTATTCCTTCACAGGCTTGTCTTATAC 143
 Db 7 GCCTGGCGCTGAGCTGCATCATGTGCGGTCTGTGGCTGGGAGGTTTCATGTCTCTGC 66
 QY 144 TGAATTTGCGAGCTCGCTCTGCCTACCAAGCTGACATGTACTTCCACCAACTGGGCCAGT 203
 Db 67 TGATGATCCCATGGGCTCTGCTGCAAAACTGGTCTGCTACTTCCCAACTGGGCCAGT 126
 QY 204 ACCGGCCAGGCTGGGGCGTTTCATGCTGACAAATCGAACATCGACCCCTGTGTACCCACC 263
 Db 127 ACAGACAGGGGAGGCTCGCTTCTGCTGCAAGGACTTGGACCCCGAGCTTTGCACCCACC 186
 QY 264 TGATCTAGCGCTTTGCTGGGAGGAGCAACACAGAGATCACCAACCATGCAATGGAAACATG 323
 Db 187 TCATCTAGCGCTTCGCTGGCATGACCAACCAAGCTGAGCAGCTGAGTGGAAATGAGC 246
 QY 324 TGACTCTCTACCAAGCTTTCAATGGCTGAAAAATAAGAACAGCAGCTGAAAACTCTCC 383
 Db 247 AGACTCTCTACCAAGGAGTTCAATGGCTGAAAGAGATGAATCCCAAGCTGAGACCTGT 306
 QY 384 TGGCCATTGGAGGCTGGAACCTTGGGAGTGGCTTCTTCACTGCCATGGTTTCTACTCTG 443
 Db 307 TAGCCATCGGAGGCTGGAATTTACAGCATCAGAAAGTTTCACAGATATGATGCCACGGCCA 366
 QY 444 AGAACCGCCAGACTTTCATCAGCTCAGTCAATTCCTGCGGAGTATGAGTTTTCAGC 503
 Db 367 ACAACCGTACAGCTTTGTCAACTCGGCGCATCAGTTTCTGCGCAATATACAGCTTTGACG 426
 QY 504 GGCTGGACTTTGACTGGGAGTACCTGCTCTCTGCTGGAGCCCTTCTCAGACAAAGCATC 563
 Db 427 GCCTTGACCTTGACTGGGAGTACCCAGAAAGCCAGGAGGAGCCCTGCCGTAGACAGGAGC 486
 QY 564 TCTTCACTGTCTGCTGCGAGAAATGCGTGAAGCTTTTGAGCAGGAGGCGCAAGCAGATCA 623
 Db 487 GCTTCAACAACCTTGTTACAGGACTTGGCCAAATGCTTCCAGCAGGAAGCCAGACCTCAG 546
 QY 624 ACAAGCCAGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 683
 Db 547 GGAAGGAACGCTTCTTCTGAGTGACGCGGTTCAGCTGGGAGACCTATGTGGATGCTG 606
 QY 684 GCTATGAGATCCCGCAACTGTACAGTACCTGGAGTACATCTCATGCTACGCTTACGACC 743
 Db 607 GATACGAGGTGGACAAATCGCCAGAAACCTGGATTGTTGCAACCTTATGGCTTACGACT 666
 QY 744 TCCATGCTCTGCGGAGGCTTACACTGGAGAGAACGCCCTCTTACAATATACCCGACTG 803
 Db 667 TCCATGCTCTTGGGAGAGGTCACGGGACATACACAGCCCTCTCTACAGAGGCAAGAG 726
 QY 804 ACACCGCCAGCAACGCTTCAATGCTGATATGTCATGTAAGTACTGTAAGGAGCAATG 863
 Db 727 AGAGTGTGAGCAGGAGCCAGCTCAACGTGGATGCTGTGTCACAGTGGCTGCGAAGG 786
 QY 864 GAGCAGCAGCTGAGAGCTCATGTTGGATTTCCTTACCTATGAGCAGACAACTTCTATCGA 923
 Db 787 GGACCCCTGCGCAGCAAGCTGATCTTGGCATGCTTACCTACGAGGCTCTCTTACACATGG 846
 QY 924 GCACCCCTCCACACTGGAATTTGGTGGCCCGCCACCTCTGTTGCTGCTGCTGCTGCTGCT 983

RESULT 9
 AAD03760
 ID AAD03760 standard; cDNA; 1656 BP.
 XX
 AC AAD03760;
 XX
 DT 19-JUN-2001 (first entry)
 XX
 XX Human chitinase cDNA from clone pMO-13B.
 XX
 XX Human; antifungal; chitinase; immunoglobulin; Ig; therapy;
 KW fungal infection; candidiasis; aspergillosis; coccidioidomycosis;
 KW blastomycosis; paracoccidioidomycosis; histoplasmosis; cryptococcosis;
 KW chromoblastomycosis; sporotrichosis; mucormycosis; dermatophytoses;
 KW clone pMO-13B; ss.
 XX
 OS Homo sapiens.
 XX
 XX
 FH Key Location/Qualifiers
 CDS 27..1427
 FT /*tag= a
 FT /*product= "Human chitinase from clone pMO-13B"
 FT sig_peptide 27..89
 FT /*tag= b
 FT mat_peptide 90..1424
 FT /*tag= c
 FT /*product= "Human mature chitinase from clone pMO-13B"
 XX
 XX WO200123430-A2.
 XX
 XX
 XX
 PD 05-APR-2001.
 XX
 XX
 PF 28-SEP-2000; 2000WO-US26960.
 XX
 XX
 PR 30-SEP-1999; 99US-0409918.
 XX
 XX (ICOS-) ICOS CORP.
 XX
 XX
 PI Allison DS, Dietsch GN, Gray PW, Shaw KD, Steiner BH;
 DR WPI; 2001-266141/27.
 XX P-PSDB; AAE00433.
 XX
 XX
 PT Novel chitinase immunoglobulin fusion product, useful for treating
 PT fungal infections and reducing the amount of a non-chitinase antifungal
 PT agent needed for the treatment -
 XX
 XX
 PS Claim 2; Page 34-36; 39pp; English.
 XX
 XX The present invention relates to a chitinase immunoglobulin (Ig) fusion
 XX product, comprising a human chitinase fused to at least a portion of an
 CC immunoglobulin chain. The fusion product is useful for treating fungal
 CC infections (mycoses) such as candidiasis, aspergillosis, blastomycosis,
 CC coccidioidomycosis, paracoccidioidomycosis, histoplasmosis, mucormycosis,
 CC cryptococcosis, chromoblastomycosis, sporotrichosis and dermatophytoses.
 CC The fusion protein is useful for reducing the amount of non-chitinase
 CC antifungal agent needed to exert an antifungal activity. The fusion
 CC protein is also useful for preparing a medicament for the prophylactic
 CC or therapeutic treatment of fungal infections. Chitinase immunoglobulin
 CC fusion product has unexpectedly improved serum half-life and formulation
 CC properties. The present sequence is human chitinase cDNA from clone
 CC pMO-13B. Chitinase enzyme degrades chitin which is a homopolymer of
 CC beta-(1,4)-linked N-acetylglucosamine residues.
 XX
 XX Sequence 1656 BP; 365 A; 497 C; 447 G; 347 T; 0 other;

XX The present sequence is that of cDNA clone MO-13B (ATCC 98078) encoding human chitinase (see AB076292). The clone was isolated from a cDNA library prepared from peripheral blood monocyte-derived macrophages using a probe based on human chitinase clone MO-218 (see AB07380). The 2 clones contain a single nucleotide difference in the coding region, causing amino acid position 81 of the mature protein to be either serine or glycine. Northern blots showed highest chitinase gene expression in lung and ovary tissues. Expression in lung is consistent with a protective role against pathogenic organisms that contain chitin. The invention provides human chitinase polynucleotides and polypeptides, and materials and methods for the recombinant production of human chitinase products, which are expected to be useful as products for treating fungal infections or for the development of such products. Human chitinase has a synergistic effect on the actions of other fungicides. It can be administered to improve the antifungal activity of a non-chitinase antifungal agent, especially amphotericin B or itraconazole, in the treatment of a fungal infection such as candidiasis, aspergillosis, coccidioidomycosis, blastomycosis, paracoccidioidomycosis, histoplasmosis, cryptococcosis, chromoblastomycosis, sporotrichosis, mucormycosis, dermatophytoses and pneumocystis infections (all claimed). In particular, the fungal infection involves *Candida*, *Aspergillus* and/or *Cryptococcus* spp.; whose growth is not effectively inhibited by contact with human chitinase alone.

Sequence 1656 BP; 365 A; 497 C; 447 G; 347 T; 0 other;

Query Match 29.7%; Score 483; DB 24; Length 1656;
 Best Local Similarity 60.2%; Pred. No. 1.8e-126;
 Matches 870; Conservative 0; Mismatches 545; Indels 30; Gaps 3;

QY 84 GACTGGTGTGCTGACATGCAACATGACAAAGCTTATTCCTCCACAGGCTTGTCTCTATAC 143
 DB 7 GCCTGCGCTGAGCTGCATCATGCTGGCTGTGGCTGGGAGGTTTCATGCTCTGC 66

QY 144 TGAATTTGACGCTCGGCTCGCTTACCAGCTGACATGCTACTTCCACCACTGGGCCAGT 203
 DB 67 TGATGATCCCATGCGGCTGCTGCTCAAACTGGTCTGCTACTTCCACCACTGGGCCAGT 126

QY 204 ACCGGCAGGCGCTGGGCGCTTCATGCTGACACATCGACCCCTGCTGTACCCACC 263
 DB 127 ACACAGAGGGGAGGCTCGCTTCTGCTCCAGGAGCTGGAGCCCGGCTTTCACCCACC 186

QY 264 TGATCTACGCTTGTGCTGGAGGCGAACAACAGAGATCACACCATCGAATGGAACGATG 323
 DB 187 TCATCTACGCTTGTGCTGGCATGACCAACACCACTGAGCACCCTGAGTGAATGACG 246

QY 324 TGATCTCTACCAAGCTTCAATGGCTGGAATAAGAAATAGACAGCAGCTGAAACTCTCC 383
 DB 247 AGACTCTCTACCAAGGAGTTCAATGGCTGGAAGAGATGAATCCCAAGCTGAAGACCTGT 306

QY 384 TGGCATTTGGAGGCTGGAACCTTGGGAGCTGCCCTTTCACTGGCATGTTTCTACTCCTG 443
 DB 307 TAGCCATCGAGGCTGGAATTTAGCACTCAGAGTTTCAAGATATGATGACCCAGGCCA 366

QY 444 AGAACCCGACAGCTTTCATCACCTCAGTATCAAAATCTCTGGCAGTATGAGTTGACG 503
 DB 367 ACAACCGTCAGACCTTTGTCAACTGGCCATCAGGTTTCTGGCGAAATACAGCTTTGACG 426

QY 504 GGCTGGACTTTGACTGGAGTACCCTGGCTCTGTTGGGAGCCCTCTCAGGACAGCATC 563
 DB 427 GCCTTGACCTTGAGTGGAGTACCCAGGAGCCAGGAGGAGCCCTGCGCTAGACAAAGAGC 486

QY 564 TCTTCACTGTCTGGTGCAGAAATGCTGAAGCTTTTGAAGCAGGAGCCCAAGCAGATCA 623
 DB 487 GCTTCAACACCTGGTACAGACTTGGCCAAATGCTTCCAGCAGGAAGCCAGACCTCAG 546

QY 624 ACAAGCCAGGCTGATGGTCACTGCTCAGTAGTGTGGCATCTCCAATATCCAGTCTG 683
 DB 547 GGAAGGAACGCTTCTTCTCAGTGCAGCGTTCAGAGCTGGGAGACCTATGTGGATGCTG 606

QY 684 GCTATGAGATCCCCCAACTGTACAGTACCTGGACTACATCCATGTGTATGATGACCTACGACC 743
 DB 607 GATACGAGTGGGACAAATCGCCAGAACCTGGATTGTTGCAACCTTTATGSCCTACGACT 666

QY 744 TCCATGCTCTGGGAGGCTACACTGGAGAGACAGCCCTCTACAAATACCCGACTG 803
 DB 667 TCCATGCTCTGGGAGGAGTACAGGAGACATACAGCCCTCTTACAAGAGGCAAGAG 726

QY 804 ACACCGCAGCAACGCTACTCAATGTGGATTATGTCAAGTACTGTGAAGACAAATG 863
 DB 727 AGATGGTGCAGAGCAGCTCAACGTGGATGCTGTGTCACACATGGCTGCAGAGG 786

QY 864 GAGACAGCTGGAAGCTCATCGTTGGATTCCCTACTATGGACACAACTTTCCTCGA 923
 DB 787 GGACCCCTGCAGCAAGCTGATCTTGGCATGCTTACCTAGGACGCTCTCTTACACATGG 846

QY 924 GCACCCCTCCACACATGGATTTGGTGGCCCACTCTGGTGTGGTCTGCTGCTGGGCGCT 983
 DB 847 CCTCCTCATCAGACACAGAGTGGGGCCCGCCAGCCAGGGTCTGGCACTCCAGGCGCT 906

QY 984 ATGCCAAGGAGTCTGGGCTTACTACGAGATCTGTACCTTCTGAAAAAATGGAG 1043
 DB 907 TCACCAAGGAGGAGGATGCTGGCTTACTATGAAGTCTGCTCT-----GGAGGGG 960

QY 1044 CCAGTACGGATGGATGCCCCCTCAGGAAGTCCCTTATGCCATFACAGGGAATGTGGG 1103
 DB 961 CCACCAACAGAGAAATCCAGGATCAGAAAGTGGCTTACATCTTCCGGAGCAACAGTGGG 1020

QY 1104 TTGCTATGACACATCAAGAGCTTCGATATTAGGCTCAATGGCTTAAAGCACACAAAT 1163
 DB 1021 TGGCTTTGATGATGTGGAGAGCTTCAAAACCAAGGTACGCTATCTGAAGCAGAAAGGAC 1080

QY 1164 TTGAGGCGCATGCTGGGCCATTTGATCTGGATGACTTCACTGGCACTTTCTGCAACC 1223
 DB 1081 TGGCGGGGCCATGCTGGCACTGGACTTAGATGACTTTCGGGCTTCTCTGCAACC 1140

QY 1224 AGGCAAGTTTCCCTTAATCTCCACCTGAAGAGGCCCTCGGGCTCGACAGTGCAAGTT 1283
 DB 1141 AGGCGCGATACCCCTCATCCAGCGCTACGCGGAGGAACAGTCTTCCATCTTGCCTT 1200

QY 1284 GCAGGCTCCAGCTCAGCCCATTTAGCCAAATACTGCTGCTCCCGAGTGGCAGGGAACG 1343
 DB 1201 CAGGCACCCAG---AGCTTGAAGTTCCAAACAGAGTACGCTTCTGAACCTGAGCA-- 1255

QY 1344 GGAGCGGAGTAGCAGCTCTGGAGAGCTCGGAGGAGCTGGATGTTCTGTCTCAGAG 1403
 DB 1256-----TGGCCCGACGCTTGCAGACACACGCTTCTGCGAGGCAAG 1296

QY 1404 CCACGGCTCTACCCCGTGGCAATTAACAGAAATGCTTCTGGCACTGCTGATGAG 1463
 DB 1297 CTGATGGGCTCTATCCCAATCCTCGGGAACGGTCCAGCTTCTACAGCTGTGACGCGGGC 1356

QY 1464 TCAGTACACAGCAACTGCGAGCGGCTGTCTTCGACACACAGCTGTGATGCTGCA 1523
 DB 1357 GGCTGTCCAGAAAGTGGCCGACAGGCTGTGTGTTCACCAACTCTCTGCAAAATGCTGCA 1416

QY 1524 ACTGG 1528
 DB 1417 CCTGG 1421

RESULT 11

AAT89181

ID AAT89181 standard; cDNA; 1768 BP.

XX AC AAT89181;

XX DT 27-APR-1998 (first entry)

XX Human chitotriosidase variant cDNA.

XX Chitotriosidase; tissue remodelling disorder; diagnosis; therapy;
 XX Chitotriosidase; atherosclerosis; human; ss.

RESULT 12

RESOLU. I.
AAT50833

AA150833
ID AAT50833 standard: cDNA: 1643 BP

ID
 YY
 AA150833

XX
AC
DATE 50833.

AC
AA150833;
yy

. 24-MAR-1997 (first entry)
Human chitinase cDNA clone chi.50.

XX Chitinase; chitotriosidase; chitin; infectious disease;
KW gene therapy; vaccine; lysosomal lipodosis; Gaucher's disease;
KW leishmaniasis; sarcoidosis; X-linked adrenoleukodystrophy;
KW multiple sclerosis; drug delivery; cosmetics; food; ss.
KW

[illegible]

OS Homo sapiens.

[illegible]

Key	Location
FH	

ET	CDS	Reaction
13	1413	

```

ET
CD3
13.1413|
/*tag= a

```

EI	sia peptide		/	13	75	-
ET			.			

```

LEF
LEF
strg_perprue
13.11.13
tag=

```

mat pentide

LT
ET
mac_dep.cue

FL
XX
/ * lag =

XX
BN
W09640940-22

PN W09640940-AZ.
YY

XX
10-DEC-1996

PD 19-DEC-1996.

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

PF	06-JUN-1996;	96WO-NL00225.
XX		
PR	07-JUN-1995;	95US-0486839.
XX		
PA	(UNAM)	UNIV AMSTERDAM.
XX		
XX	Aerts JMFJG;	
PI		
XX		
DR	WPI: 1997-118698/11.	
DR	P-PSDB; AAW08584.	
XX		
PT	New human chitinase - used to treat or prevent infection by	
PT	chitin-contg. pathogens, in diagnosis and as additives to cosmetics,	
PT	foods, implants etc.	

XX
COUNTDOWN / REPORT
T-7

Claim 2; Page 39-40; 58pp; English.

A cDNA clone (AA1508333), designated chi.50, codes for a 50 kDa human chitinase (AAW08584) that is stable to many proteases, active at pH 3-8 and up to 50 deg, and stable in the circulation. Clones chi.50 and chi.39 (see also AA150834) were isolated from a human macrophage cDNA library using as probe a partial clone obtd. using primers (see also AA150835-36) based on a chitotriosidase purified from a type 1 Gaucher disease patient. The 2 cDNA clones are the result of alternative splicing of RNA. Chitinase nucleic acid can be used for large-scale prodn. of recombinant human chitinases, or can be incorporated into a gene therapy vector to treat or prevent infection by chitin-contg. pathogens.

XX
XX

SQ Sequence 1643 BP; 364 A; 490 C; 442 G; 347 T; 0 other;

100

Query Match 29.6%; Score 481.4; DB 18

Best Local Similarity 60.3%; Pred. No. 5.1e-126;

	Mismatches	Indels	Gaps
Matches	866;	Conservative	
	0;	Mismatches	541;
		Indels	30;
		Gaps	3

100

92 CTGACTGCAACCATGACAAAGCTTATTCTCCTCACAGGCTTGTCTCTATACTGAATTG 151

Qy	212	GGCTGGGGCGCTTCATGCTGTACACATCGACCCCTGCTCTGTATCCCACTGTATCTAC	271
Db	121	GGGGAGGCTCGCTTCCTGCCCAAGAGCTTGGACCCGACCTTTTGCACCCACCTCATCTAC	180
Qy	272	GCCTTTCTGGGAGGACAACAAGAGATCACCACCATCGAATGGAAAGATGTGACTCTC	331
Db	181	GCCTTCGCTGGCATGACCAACACAGCTGAGCACCATCTGAGTGGAAATGACGAGACTCTC	240
Qy	332	TACCAGAGCTTCAATGGCCCTGAATAAAGACAGCCAGCTGTAATACTCTCTGGCCATT	391
Db	241	TACCAGAGTTCATAGGCTGAAGAAGATGAATCCCAAGCTGAAGACCTGTGTAGCCATC	300
Qy	392	GGAGGCTGGAATTCGGGAGTGCCTTTTCACCTGCCATGGTTTCTACTCTCTGAGAACCG	451
Db	301	GGAGGCTGGAATTCGGGCACTCAGAAGTTTACAGATATGGTAGCCAGGCCAACAAACCGT	360
Qy	452	CAGACTTTTCATACCTCAGTCATCAAAATTCCTGGCCAGTATGAGTTTTCAGCGGCTGGAC	511
Db	361	CAGACCTTTTCAACTCGGCCATCAGTTTCTCGCAAAATACAGCTTTGAGGSCCTTGAC	420
Qy	512	TTTGTACTGGAGATACCTTGCTCTGTGGAGGCCCTCTCTCAGACAAGCATCTCTTCACT	571
Db	421	CTTGTACTGGGATACCCAGGAAGCCAGGGAGGCCCTCCGCTAGACAAGGAGCGCTTCACA	480
Qy	572	GTCTGTGTGAGAAATCGCTGAAGCTTTTGTAGCAGGAGGCCAAGCAGATCAACAACGCC	631
Db	481	ACCTGTGTACAGACTTTGGCCATGCTTCCACGAGGAAGCCACAGCTCAGGGAAGGA	540
Qy	632	AGGCTGATGTCTACTGCTGCAGTAGTGTGGCATCTCCAATATCCAGTCTGGCTATGAG	691
Db	541	CGCCTTCTTGTAGTGCAGCGGTTCACAGTGGGAGCCCTATGTGGATGTGGATACGAG	600
Qy	692	ATCCCCAATGTTCAGTACCTGGGACTACATCCATGTCTATGACTACCACTTCCATGGC	751
Db	601	GTGGACAAAATGCCAGAAACCTTGGATTTTGTCAACCTTATGGCTTACGACTTCCATGGC	660
Qy	752	TCCTGGGAGGGCTTACACTGGAGAGACAGAGCCCTCTTACAATATCCCGACTGACACCGGC	811
Db	661	TCCTGGGAGAAGTCCAGGGACATACAGAGCCCTCTACAAGAGGAAGAGAGTGT	720
Qy	812	AGCAAGCCTTACCTCAATGTGGATATGTCTATGACTACTTGGAGGACAAATGGAGACCA	871
Db	721	GCAGCAGCCAGCTCAACGTGGATGTGTGTGTCGAACAGTGGCTGCAGAGGGGAGCCCT	780
Qy	872	GCTGTGAAGCTCATCTGTGGATTCCTTACCTATGGACAACTTCATCTGTGACCAACCC	931
Db	781	GCCAGCAAGCTGATCTTGGCATGCCTACCTACGGAGCTCTCTTCACACTGGCCTCTCA	840
Qy	932	TCCAACACTGGAATTTGGTGGCCCCCACTCTGTGTGTGTCTCTCTGGGCCCTATGCAAG	991
Db	841	TCAGACACCAAGATGGGGGGCCCCAGCACAGGGTCTGGCAGCTCCAGGCGCCTTTCACCAAG	900
Qy	992	GAGTCTGGGATCTGGCTTACTACGAGATCTTACCTTCTGAAAATGGAGCACTCTAG	1051
Db	901	GAAGGAGGATGTGCCCTACTATGAAGTCTGTCTCT-----GGAAGGGGGGCCACCAA	954
Qy	1052	GGATGGGATGCCCTCAGGAAGTGCCTTATGCTATCAGGGCAATGTGGGTGGCTAT	1111
Db	955	CAGAGAATCCAGGATCAGAAGTGCCTTACATCTTCCGGGACAAACAGTGGGTGGCTTT	1014
Qy	1112	GACAACATCAAGAGCTTCGNATTTAAGGCTCAATGGCTTAAGCACAACAAATTTGAGGC	1171
Db	1015	GATGATGTGGAGAGCTTCAAAACCAAGGTCAGCTATCTGAAGCAGAAGGACTGGCGGG	1074
Qy	1172	GCCATGCTCTGGGCCATTGATCTGGATGACTTTCACCTGGCACTTCTGTCAACAGGGCAAG	1231
Db	1075	GCCATGGCTGGGCATGTGACCTTAGATGACTTTGCCGGCTTCTCTGTCAACAGGGCCGA	1134
Qy	1232	TTTCCCTTAATCTCCACCTGAAGAGGCCCTCGGCTGCAGAGTCAAGTTGACGGGT	1291
Db	1135	TACCCCTCATCAGACGCTACGGCAGGAACCTGAGTCTTCCATACCTTGCCTCTCAGGCACC	1194

Db 924 CCT-----GGAAGGGGGCCACCACAGAGAAATCCAGGATCAGAAAGGTGCCCTACATCT 977
 QY 1086 ATCAGGGCAATGTGTGGTTGGTATGACAACTCAAGAGCTTCGATATTAAGGCTCAAT 1145
 Db 978 TCCGGACACACAGTGGTGGCTTTGATGATGTGGAGAGCTTCAAAACCAAGTCACT 1037
 QY 1146 GGCTTAAGCACAACAAATTTGGAGGGCCCATGTCTGGGCCATTTGATCTGGATGACTTCA 1205
 Db 1038 ATCTGAAGCAGAAGGACTGGCGGGGCCATGTCTGGGCACCTGGACTTAGACTTTG 1097
 QY 1206 CTGGCACTTCTGCAACCGGGAAGTTTCCCTTAATCTCCACCTGAAGAAGCCCTCG 1265
 Db 1098 CCGGCTTCTCTCAACACGGGCGGATACCCCTCATCCAGAGCTTACGGCAGGACTGA 1157
 QY 1266 GCCTGACAGTGAAGTTGCACGGCTCCAGCTCAGCCCAATGAGCCCAATTAAGTCTGCTC 1325
 Db 1158 GTCTTCATCTTGCCTTCAGGCACCCAG---AGTTGAAGTTCCTCAAAACAGGTCAGC 1214
 QY 1326 CCAAGTGGCAGCGGGAACGGAGCGGAGTAGCAGCTCTGAGGCGCTCGGGAGGCAAGTG 1385
 Db 1215 CCTCTGAACCTGAGCA-----TGGCCCCAGCCCTGGACAAGACA 1253
 QY 1386 GATTCGTGTCTCAGAGCAACGGCTCTACCCCGTGGCAATTAACAGAAATCCCTTCT 1445
 Db 1254 CGTCTCCAGGGCAAGCTGTATGGGCTCTATCCCAATCTCGGGAACGGTCCAGCTCT 1313
 QY 1446 GGCACCTCGTGAATGAGTACAGTACAGCAGAGAACTGCCAGGCGGCTTGTCTTCGACA 1505
 Db 1314 ACAGCTGTGACGGGCGGCTGTCCAGCAAGAGCTGCCGACAGGCTGTGTTCAGCA 1373
 QY 1506 CCAGCTGTGATGCTGCAACTGG 1528
 Db 1374 ACTCTGCAAAATGCTGCACCTGG 1396

RESULT 14

AAZ21847
 ID AAZ21847 standard; DNA; 1636 BP.

XX AAZ21847;

XX 10-DEC-1999 (first entry)

XX MO-218 clone of human Chitinase, with noncoding 5'/3' regions.

XX chitin; fungal infection; immunocompromised; AIDS; chemotherapy;

XX organ transplant; parasite; chitin-binding; allele; vector;

XX truncated protein; ds.

XX Homo sapiens.

OS Key Location/Qualifiers

XX CDS 2..1402

FT /tag= a

FT /product= Human_Chitinase

FT 2..65

FT /tag= b

FT /product= Signal_peptide

FT 66..1402

FT /tag= c

FT /product= Mature_protein

XX WO946390-A1.

XX 16-SEP-1999.

XX 12-MAR-1999; 99WO-US05343.

XX 12-MAR-1998; 98US-0039198.

XX (ICOS-) ICOS CORP.

XX Gray PW, Tjoelker LW;

XX

PI

XX

DR WPI; 1999-551417/46.

DR P-PSDB; AAY42425.

XX Novel chitin-binding fragments of human chitinase used to treat fungal

PT infections in animals

PT

PS Example 1; Page 55-57; 83pp; English.

XX

XX This is the nucleotide sequence of an allelic form of the human

CC chitinase enzyme, which is capable of degrading Chitin (a linear

CC homo polymer of beta-1,4-linked N-acetylglucosamine residues).

CC Chitinase fragments can be used to screen for proteins or other

CC molecules that specifically bind to the chitin-binding domain of human

CC chitinase or that modulate its activity. These compounds are useful for

CC immunization, as well as for purifying chitinase, as well as for

CC detection and quantification of chitinase. Polynucleotide fragments of

CC the invention are useful as a source of probes and primers, and to

CC express the proteins recombinantly. The chitinase fragments, when

CC conjugated to antifungal compounds, are used to treat animals,

CC especially humans, infected with chitin-containing parasites such as

CC fungi. Fungal infection treated include candidiasis, aspergillosis,

CC coccidioidomycosis, blastomycosis, paracoccidioidomycosis,

CC mucormycosis, histoplasmosis, cryptococcosis, chromoblastomycosis,

CC Chitin can be degraded by the enzyme chitinase. Use of whole chitinase

CC protein for treating infections, especially fungal infections, is

CC problematic. In view of the increasing incidents of life-threatening

CC fungal infection in e.g. immunocompromised individuals, there exists a

CC need for identifying new compounds for treating fungal infection. The

CC chitin-binding fragments of the present invention provide this need.

XX

SQ Sequence 1636 BP; 361 A; 491 C; 440 G; 344 T; 0 other;

Query Match

Best Local Similarity

Matches

Conservative

Mismatches

Indels

Gaps

Length

1636;

29.68; Score 481; DB 20;

60.88; Pred. No. 6.6e-126;

0; Mismatches 520;

30; Gaps 3;

126 CAGGTCTTGTCTTATCTATCAATTTGCAGCTCGCTCTGCTACCAAGCTGACATGCTACT 185

24 CAGGTTTCATGGTCTCTGCTGATGATCCATGGGGCTCTGCTGCANAACTGGTCTGCTACT 83

186 TCACCAACTGGGGCCAGTACCGGCCAGGCGCTGGGGGCGCTTCATCGCTGACAAATGACAC 245

84 TCACCAACTGGGGCCAGTACAGACAGAGGGGAGGCTGCTTCTCTGCCCAAGGACTTGGACC 143

246 CCGTCTCTGTACCACTGATCTACGCTTTCTGGGAGGAGAGACAGACAGATGACCA 305

144 CAGGCTTTGCAACCCACCTCATCTACGCTTCTGCTGCAATGACCAACCAAGCTGAGCA 203

306 CCATCGAATGGAACGATGTGACTCTCTACCAAGCTTTCAATGGCTGAAATAAGAACCA 365

204 CCACTGAGTGAATGACGAGACTCTCTACCAAGGAGTCAATGGCTGAAAGATGAATC 263

366 GCAGCTGAAACTCTCTCGGCCATTTGGAGGCTGGAACCTTCGGGACTGCCCTTTCAGTG 425

264 CCAAGCTGAAGACCTCTTACGCTAGGAGGCTGGAATTTCCGCACTCAGAACTTCACAG 323

426 CCATGTTTCTACTCTCTGAGAACCGCCAGACTTTTCATCCTCAGTCATCAATTCCTGC 485

324 ATATGGTAGCCAGGCCAACACCGTCAGACCTTTGTCAACTCGGCCATCAGGTTTCTGC 383

486 GCCAGTATGAGTTTACGGGCTGGACTTTGACTGGGAGTACCTTGGCTCTCGTGGGAGCC 545

384 GCAATACAGCTTTGACGGGCTTTGACCTTGACTGGGAGTACCCAGGAAGCCAGGGAGCC 443

546 CTCCTCAGGACAGCATCTCTTCACTGTCTGTGTGAGGAAATTCGCTGAAGCTTTTGAGC 605

444 CTGCCGTAGACAAGAGGCGCTTCAACAACCTGTGTACAGGACTTTGGCCAATTCCTTCAGC 503

606 AGGAGGCCAAGCAGATCAACAAGCCCGGCTGTATGCTACTGCTGAGTAGCTGCTGGCA 665

504 AGGAAGCCACAGACCTCAGGGAAGGACAGCCCTTCTTCTGAGTCGACGGGTTCCAGCTGGC 563

```
KW 666 TCTCAATATCCAGTCTGGCTATGAGATCCCAACTGTACAGTACCTGGACTACATCC 725
  || || || || || || || || || || || || || || || || || || || || ||
KW 564 AGACCTATCTGGATCTGATGAGAGGTGGAGAAATCGCCAGAACCTGGATTGTCA 623
  || || || || || || || || || || || || || || || || || || || || ||
KW 726 ATGTGATGACCTACGACCTCCATGGCTCTGGAGGGGTACACTGGAGAGAACAGCCGCC 785
  || || || || || || || || || || || || || || || || || || || || ||
DB 624 ACCTTATGGCTTACACTGCTTCCATGGCTCTGGAGAAAGTCCAGGGACATACAGCCGCC 683
  || || || || || || || || || || || || || || || || || || || || ||
QY 786 TCTACAAATACCGGACTGACACCGGACGAGCAAGCCCTACCTCAATGGATTATGTCATGA 845
  || || || || || || || || || || || || || || || || || || || || ||
DB 684 TCTACAAAGAGGCAAGAGAGTGTGTCAGACGAGCCCTCAACCTGGATGCTGTGTC 743
  || || || || || || || || || || || || || || || || || || || || ||
QY 846 ACTACTGGAAGACAAATGAGCAGCAGCTGAGAAGCTCATCTGTGATTCCTTACCTATG 905
  || || || || || || || || || || || || || || || || || || || || ||
DB 744 AACAGTGGCTGAGAGGGGACCCCTGCGCAGCAGCTGATCTTGGCATGCCCTACCTAGC 803
  || || || || || || || || || || || || || || || || || || || || ||
QY 906 GACACAACTTCTCTGAGCAACCCCTCCAACTGGAATTTGGTGGCCGCCACCTCTGTGTG 965
  || || || || || || || || || || || || || || || || || || || || ||
DB 804 GAGCTCTCTTCACTGCTGCTCTCATCAGACACCAAGAGTGGGGGCCCCAGCACAGGT 863
  || || || || || || || || || || || || || || || || || || || || ||
QY 966 CTGTCTCTGCTGGCCCTATGCAAGGCTCTGGGATCTGGGCTTACTACGAGATCTGTA 1025
  || || || || || || || || || || || || || || || || || || || || ||
DB 864 CTGGCACTCCAGGCCCTTCAACAAGGAGGAGGATGCTGGCTTACTATGAGTCTGCT 923
  || || || || || || || || || || || || || || || || || || || || ||
QY 1026 CCTTCTGAAAAATGAGGCACTCAGGGATGGGATGCCCTCAGGAAGTGCCTTATGCT 1085
  || || || || || || || || || || || || || || || || || || || || ||
DB 924 CCT-----GGAAGGGGCCCAACACAGAGATCCAGGATCAGAGGTGCCCTACATCT 977
  || || || || || || || || || || || || || || || || || || || || ||
QY 1086 ATCAGGGCAATGTGGGTGGCTATGACAACTCAAGAGCTTCGATATTAAGGCTCAAT 1145
  || || || || || || || || || || || || || || || || || || || || ||
DB 978 TCCGGACACCAACGAGTGGGTGGCTTTGATGATGTGGAGAGCTTCAAAACCAAGGTCAGCT 1037
  || || || || || || || || || || || || || || || || || || || || ||
QY 1146 GCTTAAGCACACAAATTTGAGGGCCCATGCTGCGCCATGATCTGGATGACTTCA 1205
  || || || || || || || || || || || || || || || || || || || || ||
DB 1038 ATCTGAAGCAGAGGAGCTGGGGGGCCATGGTCTGGGCACTGGACTTAGATGACTTTG 1097
  || || || || || || || || || || || || || || || || || || || || ||
QY 1206 CTGGCACTTTCTGCAACGAGGCAAGTTTCCCTTAATCTCCACCTGAAGAGGCCCTCG 1265
  || || || || || || || || || || || || || || || || || || || || ||
DB 1098 CCGGCTTCTCTGCAACGAGGCGGATACCCCTCATCCAGAGCTACGCGAGGACTGA 1157
  || || || || || || || || || || || || || || || || || || || || ||
QY 1266 GCTCGAGAGTCAAGTTCGACGGCTCCAGCTCAGCCCATTTAGCCAAATTAATGTGCTC 1325
  || || || || || || || || || || || || || || || || || || || || ||
DB 1158 GTCTTCCATPACTTGGCTTCAGCACCCAG---AGCTTGAAGTTCCAAAACCAAGGTCAGC 1214
  || || || || || || || || || || || || || || || || || || || || ||
QY 1326 CCAAGTGGCAGCGGACGAGCGGAGTAGCAGCTCTGGAGGAGCTCGGAGGAGCAGTG 1385
  || || || || || || || || || || || || || || || || || || || || ||
DB 1215 CCTCTGAACCTGAGCA-----TGGCCCCAGCCCTGGACAAAGACA 1253
  || || || || || || || || || || || || || || || || || || || || ||
QY 1386 GATCTGTGCTCTGAGGCAACGCGCTCTACCCGTTGGCAATAACAGAAATGCTTCT 1445
  || || || || || || || || || || || || || || || || || || || || ||
DB 1254 CGTCTGCCAGGCAAGCTGATGGCTCTATCCCAATCCTCGGGAACGGTCCAGCTTCT 1313
  || || || || || || || || || || || || || || || || || || || || ||
QY 1446 GGCATGCTGAATGAGTACGTCACGACGAGCAATGCGAGCGCGGGCTTGTCTTGACA 1505
  || || || || || || || || || || || || || || || || || || || || ||
DB 1314 ACAGCTGTGACGCGGGGGCTTCCAGAAAGTGGCCGACAGGCTGGTGTTCAGCA 1373
  || || || || || || || || || || || || || || || || || || || || ||
QY 1506 CCAGCTGTGATGCTGCAACTGS 1528
  || || || || || || || || || || || || || || || || || || || || ||
DB 1374 ACTCTGCAATGCTGCACCTGG 1396
  || || || || || || || || || || || || || || || || || || || || ||
```

RESULT 15

AAD03759

ID AAD03759 standard; cDNA; 1636 BP.

XX

AC AAD03759;

XX

DT 19-JUN-2001 (first entry)

XX

DE Human chitinase cDNA from clone pMO-218.

XX

KW Human; antifungal; chitinase; immunoglobulin; Ig; therapy;

Db 204 CCACTGAGTGAATGACGAGACTCTCTACAGGAGTTCAATGGCCTGAAGAAGATGAATC 263
Qy 366 GCCAGCTGAACACTCTCTGCGCATTTGGAGCTGGAACCTTCGGACACTGCCCTTTCACTG 425
Db 264 CCAAGCTGAAGACCCCTGTAGCCATCGGAGCTGGAATTTGGCAGCTCAGAGTTCCACAG 323
Qy 426 CCATGGTTTCTACCTCTGAGAACCGCCAGAGCTTTATCATCCTCAGTCAATCAATCTCTGC 485
Db 324 ATATGGTAGCCACCGCCCAACACCGCTCAGACCTTTGTCAACTCGGCCATCAGGTTTCTGC 383
Qy 486 GCCAGTATGAGTTGAGCGGCTGAGCTTTGACTGGGAGTACCTTGGCTCTCGTGGAGCC 545
Db 384 GCAATACAGCTTTGAGCGGCTTGACCTTGAGTGGAGTACCCAGGAAGCGAGGAGCC 443
Qy 546 CTCCTCAGGCAACGACTCTTTCACCTGTCTGGTGCAGGAATCGCTGAAGCTTTTGAGC 605
Db 444 CTGCGGTAGACAAGAGCGCTTTCACACCCCTGGTACAGGACTTGGCCAAATGCCCTTCCAGC 503
Qy 606 AGGAGGCCAGAGACTCAACAGCCCGAGGCTGATGGTCACTGCTGCACTAGCTCTCTGSCA 665
Db 504 AGGAAGCCAGACCTCAGGAAGGAAGACGCTTCTCTGAGTGCAGCGGTTCCAGCTGGGC 563
Qy 666 TCTCCAATATCCAGTCTGGCTATGAGATCCCCCAACTGTACAGTACTCTGAGCTACATCC 725
Db 564 AGCCTATGTGGATGCTGGATACGAGGTGGACAAATCGCCCAAGACCTGGATTTGTCA 623
Qy 726 ATGTCACTGACCTAGCAGCTCCATGGCTCTGGGAGGCTACACTGGAGAGAACAGCCGCC 785
Db 624 ACCTATGGCCTAGGACTTCCATGGCTTCTGGGAGAGGTCACGGGACATAACAGCCGCC 683
Qy 786 TCTACAAATACCCGACTGACACCGGCGAGCAACGCTTACCTCAATGTGGATTTATGTCATGA 845
Db 684 TCTACAAAGAGCAAGAGAGAGTGGTGCAGCAGCCGCTCAAGCTGGATGCTGCTGTC 743
Qy 846 ACTACTGGAAGGCAATGGAGCACCAGCTCAGAGCTCATCTGTTGGATTCCCTACCTATG 905
Db 744 AACAGTGGCTGCAAGAGGGACCCCTGCCAGCAGCTGATCCTTGGCATGCCCTACCTACG 803
Qy 906 GACACAACTTCATCCTCAGCAACCCCTCCAAACACTGGAATTTGGTCCGCCACCTCTGTTG 965
Db 804 GAGCCTCTTCACACTGGCCTCCTCATCAGACACCAGAGTGGGGCCGCCAGCCACAGGT 863
Qy 966 CTGCTCTCTGCGCCCTATGCCAAGGAGTCTGGGATCTGGGCTTACTACGAGATCTGTA 1025
Db 864 CTGSCACTCCAGGCCCTTCCACCAAGAGGAGGATGCTGGCCTACTATGAAGTCTGCT 923
Qy 1026 CTTCTCTGAAATGAGGACACTCAGGATGGGATGCCCTCAGGAAGTGCCTTATGCCCT 1085
Db 924 CCT-----GGAAGGGGCCCAACACAGAGATCCAGGATCAGAGGTGCCCTACATCT 977
Qy 1086 ATCAGGGCAATGTGGGTTGGCTATGACAACATCAAGAGCTTCGATATTAAAGGCTCAAT 1145
Db 978 TCCGGGACACACCGAGTGGGCTTTGATGATGTGGAGAGCTTCAAAACCAAGGTCACT 1037
Qy 1146 GGCTTAAGCACACAAATTTGGAGGCCCATGCTGCGGCAATGATCTGGATGACTTCA 1205
Db 1038 ATCTGAAGCAGAGGAGTGGGGGGCCATGCTGGGCACTGGACTTAGATGACTTG 1097
Qy 1206 CTGSCACTTCTGCACCGGCAAGTTTCCCTTAATCTCCACCCTGAAGAAGGCCCTCG 1265
Db 1098 CCGGCTTCTCTGCAACGAGGGCGGATACCCCTCATCCAGAGCTACGCGGAGAACTGA 1157
Qy 1266 GCCTGAGAGTGAAGTTGCACGGCTCCAGCTCAGCCCATTTGAGCCCAATAACTGCTGCTC 1325
Db 1158 GTCTTCCATACTTGCCTTCAGGCACCCAG--AGCTTGAAGTTCCAAACACAGGTCAGC 1214
Qy 1326 CCAAGTGGCAGGGAGCGGAGGAGTAGAGCTCTGAGGAGAGCTCGGGAGGAGTG 1385
Db 1215 CCTCTGAACCTGAGCA-----TGCCGCCAGCCCTGGACAAGACA 1253
Qy 1386 GATTCGTGCTGTACAGGCAACGGCTCTACCCCGTGGCAATAACAGAAATGCCCTTCT 1445
Db 1254 CGTTCGCCAGGGCAAGCTGATGGGCTCTATCCCAATCCTCGGGAACGCTCCAGCTTCT 1313

. . .

Qy 1446 GGCACCTGCTGAATGGAGTCACTACCAAGCAAGAACTGCCAGGCGGGCTTGTCTTCGACA 1505
Db 1314 ACAGCTGTGCAGCGGGGGCTGTTCCAGCAAAAGTGGCCGACAGGCTGTGTTCAGCA 1373
Qy 1506 CCAGCTGTGATGTGCTGCAACTGG 1528
Db 1374 ACTCCTGCAAAATGCTGCACCTGG 1396

Search completed: July 3, 2003, 05:03:18
Job time : 377.46 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2003, 04:38:16 ; Search time 78.4127 seconds
(without alignments)
6355.472 Million cell updates/sec

Title: US-10-004-219B-2
Perfect score: 1625
Sequence: 1 gcttcagtcggtgtga.....ctgcaataaaatcagcagtc 1625

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCRUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	483	29.7	1656	US-09-039-198A-3	Sequence 3, Appl
2	483	29.7	1656	US-08-877-599-3	Sequence 3, Appl
3	483	29.7	1656	US-09-267-574-3	Sequence 3, Appl
4	481.4	29.6	1643	US-08-486-839-3	Sequence 3, Appl
5	481.4	29.6	1643	US-09-151-011-3	Sequence 3, Appl
6	481.4	29.6	1643	US-09-343-623-3	Sequence 3, Appl
7	481	29.6	1636	US-09-039-198A-1	Sequence 1, Appl
8	481	29.6	1636	US-08-877-599-1	Sequence 1, Appl
9	481	29.6	1636	US-09-267-574-1	Sequence 1, Appl
10	477.2	29.4	1713	US-08-486-839-5	Sequence 5, Appl
11	477.2	29.4	1713	US-09-151-011-5	Sequence 5, Appl
12	477.2	29.4	1713	US-09-343-623-5	Sequence 5, Appl
13	369	22.7	1433	US-08-694-915-1	Sequence 1, Appl
14	354.8	21.8	1526	US-08-694-915-3	Sequence 3, Appl
15	339.8	20.9	1681	PCT-US94-07754-4	Sequence 4, Appl
16	338.2	20.8	1681	US-08-581-527-4	Sequence 4, Appl
17	156	9.6	2452	US-08-524-051-1	Sequence 1, Appl
18	127.2	7.8	1677	US-09-545-814-13	Sequence 13, Appl
19	127.2	7.8	1677	US-09-545-814-15	Sequence 15, Appl
20	127.2	7.8	1749	US-09-545-814-4	Sequence 4, Appl
21	127.2	7.8	1749	US-09-545-814-6	Sequence 6, Appl
22	127.2	7.8	1919	US-09-545-814-31	Sequence 31, Appl
23	127.2	7.8	1919	US-09-545-814-33	Sequence 33, Appl
24	127.2	7.8	2610	US-09-545-814-1	Sequence 1, Appl
25	127.2	7.8	2610	US-09-545-814-3	Sequence 3, Appl
26	120.6	7.4	1478	US-09-545-814-28	Sequence 28, Appl
27	120.6	7.4	1478	US-09-545-814-30	Sequence 30, Appl

28	111.8	6.9	1608	4	US-09-292-225-20	Sequence 20, Appl
29	111.8	6.9	1608	4	US-09-292-225-22	Sequence 22, Appl
30	111.8	6.9	1665	4	US-09-292-225-17	Sequence 17, Appl
31	111.8	6.9	1685	4	US-09-292-225-19	Sequence 19, Appl
32	111.8	6.9	1752	4	US-09-292-225-14	Sequence 14, Appl
33	111.8	6.9	1752	4	US-09-292-225-16	Sequence 16, Appl
34	103	6.3	1470	4	US-09-292-225-40	Sequence 40, Appl
35	103	6.3	1470	4	US-09-292-225-42	Sequence 42, Appl
36	103	6.3	1527	4	US-09-292-225-37	Sequence 37, Appl
37	103	6.3	1527	4	US-09-292-225-39	Sequence 39, Appl
38	103	6.3	1621	4	US-09-292-225-34	Sequence 34, Appl
39	103	6.3	1621	4	US-09-292-225-36	Sequence 36, Appl
40	68.8	4.2	1167	1	US-07-939-501A-6	Sequence 6, Appl
41	68.8	4.2	1167	4	US-08-448-398-10	Sequence 10, Appl
42	68.8	4.2	1320	1	US-07-939-501A-14	Sequence 14, Appl
43	68.8	4.2	1364	1	US-07-939-501A-13	Sequence 13, Appl
44	68.8	4.2	1405	1	US-07-939-501A-9	Sequence 9, Appl
45	68.8	4.2	1701	1	US-07-939-501A-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1
US-09-039-198A-3
Sequence 3, Application US/09039198A
Patent No. 6200951
GENERAL INFORMATION:
APPLICANT: Gray, Patrick W.
TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive/6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/039,198A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 27866/34391
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1656 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 27..1424
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 90..1424
US-09-039-198A-3

Query Match 29.7%; Score 483; DB 4; Length 1656;
Best Local Similarity 60.2%; Pred. No. 4.6e-132;
Matches 870; Conservative 0; Mismatches 545; Indels 30; Gaps 3;

Qy	144	TGAATTTGCAGCTCGGGCTCTGCCTACACGCTGCACATGCTACTTACCAACTGGGCCCAGT	203
Db	67	TGATGATCCCATGGGGCTCTGCTGCAAAACATGGTCTGCTACTTTCACCAACTTGGSCCAGT	126
Qy	204	ACGGCCAGCGCTGGGGCGCTTCATGCGCTGACACATCGACCCCTGCTCTGTACCCACC	263
Db	127	ACAGACAGGGGAGGCTCGCTTCTCGCCCAAGGACTTGGACCCCGAGCTTTGCACCCACC	186
Qy	264	TGATCTACGCGCTTTGCTTGGCAGGCAGAAACACAGATCACACCATCGAATGGAACGATG	323
Db	187	TCATCTACGCGCTTCGCTGGCATGACCAACACACAGCTGAGCACCACCTGAGTGGATGACG	246
Qy	324	TGACTCTCTACCAAGCTTTCAATGGCGTGAATAAAGAACAGCAGCTGAAAACCTCTCC	383
Db	247	AGACTCTCTACCAAGGAGTTCAATGGCGTGAAGAAGATGAATCCCAAGCTGAAGACCCCTGT	306
Qy	384	TGGCCATTGAGGCTGGAACCTTGGGACTTCGCCCTTTCACCTGCCATGGTTTCTACTCCTG	443
Db	307	TAGCCATTCGGAGGCTTGAATTTTCAGCACTCAGAACTTCACAGATATGGTAGCCACGGCCA	366
Qy	444	AGAACGCCAGACTTTCATCACCTCAGTCAATCAAAATCCCTGGCGCAGTATGAGTTTCAGC	503
Db	367	ACAAOCCTCAGACCTTTGCTCACTTGGCCATCAGGTTTCTGGCGAAATCAGCTTTGACG	426
Qy	504	GGCTGGACTTTGACTGGGAGTACCTTGGCTCTCGTGGGAGCCCTCTCAGGACAAAGCATC	563
Db	427	GCCTTGACCTTGACTGGGAGTACCCAGGAGCAGGGGAGCCCTGCCGTAGACAAAGGAGC	486
Qy	564	TCCTTACTGTCCTGGTGCAGGAATGGGTGAAGCTTTTGAGCAGGAGGGCCAAAGCATCA	623
Db	487	GCCTTCAACACCCCTGGTACAGAGCTTGGCCAAATGCCCTTCAGCAGGAAGCCACAGCTCAG	546
Qy	624	ACAAGCCAGGCTGATGGTCACTGCTGCAGTAGTCTGTGGCATCTCCAAATATCCAGTCTG	683
Db	547	GGAGGAACGCCCTCTCTGAGTGCACGGGTTCCAGCTGGCGACACTATGTGGATGCTG	606
Qy	684	GCTATGAGATCCCCAACTGTCACAGTACCTGAGCTACATCCATGTGTCATGACCTTAGGAC	743
Db	607	GATACGAGGTGGACAAAATGCCCGACAACTGGATTTGTCAACCTTATGGCTTACGACT	666
Qy	744	TCATGCTCCTCGGAGGGCTACACTGGAGAGACAGCCCCCTCTACAAATACCCGACTG	803
Db	667	TCATGCTCTTGGGAGAAGGTCACGGGACATACAGCCCCCTCTACAGAGGCAAGAAG	726
Qy	804	ACACCGCAGCAACGCCCTACTCAATGTGGATTATGTGATGAATCTACTGGAAGGACAAATG	863
Db	727	AGAGTGTGCAGCAGCCAGCCTCAACGTGGATGCTGCTGTGCACAGTGGCTGCAGNAGG	786
Qy	864	GAGCACCAGCTGAGAAGCTCATCGTTGGATTCCCTACCTATGGACACAACTTCATCCTGA	923
Db	787	GGACCCCTGCCACCAAGCTGATCCTTGGCATGCCCTACCTACGGACGGCTCCTTTCACACTGG	846
Qy	924	GCAACCCCTCCACACTGGAAATTTGGTGGCCCCCACCCTCTGGTGTGGTCTGCTGGGCCCT	983
Db	847	CTCTCTATCAGACACAGAGTGGGGGCCCCACAGGCTCTGGCAGCTTCCAGGGCCCT	906
Qy	984	ATGCCAAGGAGTCTGGGATCTGGGCTTACTTACGAGATCTGTACTTCTTCTGAAAATGGAG	1043
Db	907	TCACCAAGGAGGAGGATGCTGGCCTACTATGAAGTCTGCTCCT-----GGAAGGGG	960
Qy	1044	CCACTCAGGATGGGATGCCCTCAGGAAGTGCTTTATGGCTATCAGGGCAATGTGTGGG	1103
Db	961	CCACCAACAGACAAATCCAGGATCAGAAGGTGCCCTACATCTTCCGGGACCAACGATGGG	1020
Qy	1104	TTGGCTATGACACATCAGAGGCTTCGATATTAAAGGCTCAATGGCTTAAAGCACAAAT	1163
Db	1021	TGGGCTTTGATGATGTGGAGAGCTTCAAAACCAAGGTTCAGCTATCTGAAGCAGAAAGGAC	1080
Qy	1164	TTGAGGCGCATGGTCTGGGCCATTGATCTGATGACTTCACTGGCAGCTTTCTGGAACC	1223
Db	1081	TGGGCGGGCCATGGTCTGGGCACTGGACTTAGATGACTTTTGGCGGCTTCTCTCTGAAACC	1140
Qy	1224	AGGCAAGTTTCCCTTAACTCCACCCCTGAAGAAGGCCCTCGGCCCTGCAGAGTGCAGTT	1283

1141	AGGGCGGATACCCCTCATCCAGACGGTACGGCAGGAACCTGAGTCTTCCATCTTGCCCTT	1200
1284	GCAGGGCTCCAGCTCAGCCCATTTGAGCCAAATAACTGCTGCTCCAGTGGCAGCGGGAACG	1343
1201	CAGGCACCCGAG---AGCTTGAAGTTCCAAAACAGGTACGCCCTCTGAACCTGAGCA--	1255
1344	GGAGCGGGAGTAGCAGCTCTGGAGGCAGCTCGGGAGGCAGTGGATTCTGTGCTGTGACAG	1403
1256	-----TGGCCCCAGCCCTGGACAAGACACGTTCTGCCAGGGCAAG	1296
1404	CCAAAGGCCTCTACCCGGTGGCAAAATACAGAAATGCCCTCTGGCACTCGGTCGAATGGAG	1463
1297	CTGATGGGCTCTATCCCAATCTCTCGGAAACGGTCCAGCTTCTACAGCTGTGACAGCGGGC	1356
1464	TCACGTACCAAGCAGAACTGCCAGCGGGCTTCTTTCGACACACAGCTGTGATTGCTGCA	1523
1357	GGCTGTTCAGCAAAAGCTGCCGACAGGCCCTGGTGTTCAGCAACTCCTGCAAAATGCTGCA	1416
1524	ACTGG	1528
1417	CCTGG	1421

RESULT 4
 US-08-486-839-3
 ; Sequence 3, Application US/08486839
 ; Patent No. 5928928
 ; GENERAL INFORMATION:
 ; APPLICANT:
 ; TITLE OF INVENTION: A human chitinase, its recombinant
 ; production, its use for decomposing chitin, its use
 ; TITLE OF INVENTION: in therapy or prophylaxis against infection diseases.
 ; NUMBER OF SEQUENCES: 16
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Hoffmann & Baron
 ; STREET: 350 Jericho Turnpike
 ; CITY: Jericho
 ; STATE: New York
 ; COUNTRY: United States of America
 ; ZIP: 11758
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; SOFTWARE:
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/486,839
 ; FILING DATE: 07 - June - 1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Baron, Ronald J.
 ; REGISTRATION NUMBER: 29,281
 ; REFERENCE/DOCKET NUMBER: 294-26
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (516) 822-3550
 ; TELEFAX: (516) 822-3582
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1643 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: unknown
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: cDNA
 ; HYPOTHEetical: NO
 ; US-08-486-839-3

Query Match 29.6%; Score 481.4; DB 2; Length 1643;
Best Local Similarity 60.3%; Pred. No. 1.3e-131;
Matches 866; Conservative 0; Mismatches 541; Indels 30; Gaps 3;
QY 92 CTGACTGCACCATGACAAGCTATTCTCTCTCACAGGCTCTTGTCTCTTATCTGTAATTG 151
||||| |

QY	152	CAGCTGGGCTCTGCCTACACAGCTGACATGCTACTTCCAACTTGGGCCCCAGTACCGGCCA	211
Db			
Db	61	CCATGGGGCTGCTGCCAAACCTGGCTGCTACTTACCACCTGGGCCCCAGTACAGACAG	120
QY	212	GGCCTGGGGGCTTCATGCTGTGACACATCGACCCCTGCTCTGTATCCCACTGATCTAC	271
Db			
Db	121	GGGAGGCTCGCTTCTGCGCCCAAGGACTTGGACCCAGGCTTGGACCCCACTCATCTAC	180
QY	272	GCCTTTGCTGGGAGGAGAAACAGAGATCACCACCATCGAATGGAAAGCAATGTCATCTC	331
Db			
Db	181	GCCTTGCCTGGCATGACCAACCACAGCTGAGCACCACCTGAGTGGGAATGACGAGACTC	240
QY	332	TACCAAGCTTTCATGGCCTGAAAAATAGAACAGCCAGCTGAAAACTCTCTCTGGGCATT	391
Db			
Db	241	TACCAGAGAGTTCAATGGCCTGAGAAGATGAATCCCAAGCTGAAGACCCCTGTTAGGCATC	300
QY	392	GGAGGCTGGAACTTCGGGAGCTGCCCTTTTCACCTGCCATGGTTTCTACTCTCTGAGAACCGC	451
Db			
Db	301	GGAGGCTGGAATTCGGCACTCAGAAGTTCACAGATATGGTAGCCAGGCCAACACACCGT	360
QY	452	CAGACTTTCATCACTCACTGATCATCAATTCCTCGGCCAGTATGAGTTTGACGGCTGGAC	511
Db			
Db	361	CAGACCTTTGTCAACTCGGCCATCAGGTTTCTGCGCAATACAGCTTTGACGCCCTTGAC	420
QY	512	TTTGACTGGGAGTACCTGCTCTCGTGGGAGCCCTCCTCAGGCACAGCATCTCTTCACT	571
Db			
Db	421	CTTGACTGGGAGTACCCAGGAAGCCAGGGGAGCCCTCGCTAGACAAAGGAGCGCTTCACA	480
QY	572	GTCTGTGTGAGAAATTCGTGAAGCTTTTGACGAGGAGGCCAAGCAGATCAACAACGCC	631
Db			
Db	481	ACCCTGGTACAGGACTTGGCCATGCTCTCCAGCAGGAAGCCACAGACCTCAGGGAGGAA	540
QY	632	AGGCTGATGTCTACTGCTGCAGTAGTGTCTGGGATCTTCCAATATCCAGTCTGGCTATGAG	691
Db			
Db	541	CGCCTTCTCTGAGTGCAGCGGTTCCAGCTGGGCAGACCTATGTGGATGTGGATACGAG	600
QY	692	ATCCCCAACTGTACAGTACCTGGGACTACATCCATCTCATGACCTACGACCTCCATGGC	751
Db			
Db	601	GTGGCAAAATCGCCCAAGACCTGGATTTGTCAACCTATGCGCTACGACTTCCCATGGC	660
QY	752	TCCTGGGAGGGCTACACTGGAGAGACAGAGCCCTCTTACAAATACCCGACTGACACCGGC	811
Db			
Db	661	TCCTGGGAGAGGTCACGGGACATAACAGAGCCCTCTACAAGAGGCAAGAGAGTGCT	720
QY	812	AGCAAGCCCTACCTCAATGTGGATTATGTGATGAATCTCTGGAAGGACAATGAGGACCA	871
Db			
Db	721	GCAGCAGCCAGCCTCAACGTGGATGTGCTGTGCAACAGTGGCTGCAAGAGGGGAGCCCT	780
QY	872	GCTGAGAAGCTCATGTTGGATTCCCTTACCTATGGACACAACTTCATCTGTGACACACCC	931
Db			
Db	781	GCACGAACTGATCCTTGGCATGGCTTACCTACGAGCGTCTCTCACATGGGCTCTCTCA	840
QY	932	TCCAACATCGGAATGGTGGCCCCACCTCTGTGTGCTGGTCTGCTGGGCCCTATGCCAAG	991
Db			
Db	841	TCAGACACCAAGTGGGGGGCCCCAGGCCAGGCTCTGGCACTCCAGGCCCTTTCACCAAG	900
QY	992	GAGTCTGGGATCTGGGCTTACTAGGAGATCTGTACCTTCTGAAAAATGAGGCCACTCAG	1051
Db			
Db	901	GAAGGAGGATGTGGCCTACTATGAAGTCTGTCTCT-----GGAAGGGGGGCCACCAA	954
QY	1052	GGATGGGATGCCCTCAGGAAGTGCCTTATGCGCTATCAGGGGCAATGTGGGTTGGCTAT	1111
Db			
Db	955	CAGAAATCAGGATCAGAAGGTGCCCTACATCTTCGGGCAACAACCGTGGGTGGGCTTT	1014
QY	1112	GACAACATCAAGAGCTTCGATATTAAGGCTCAATGGCTTAAGCACACAATAATTGGAGGC	1171
Db			
Db	1015	GATGATGTGAGAGCTTCAAAACAAGGTCAGCTATCTGAAGCAGAAAGGACTGGCGGG	1074
QY	1172	GCAATGGTCTGGGCCATTGATCTGGATGACTTCACTGGCACTTTCTTGCAACACAGGCCAAG	1231
Db			
Db	1075	GCAATGGTCTGGGCACTGGATTAGATGACTTTGCGGGCTTCTCCCTGCAACACAGGCCGA	1134

1232	TTTTCCCTTAATCTCCACCCCTGAAGAGGCCCTCGGCCTGCAGAGTCAAGTTCCACGGCT	1291
Qy		
1135	TACCCCTCATTCAGACGCTACGGCAGGAACCTAGTCTTCCATPACTTGGCTTCAGGCACC	1194
Db		
1292	CCAGCTCAGCCCATTTAGGCCAATAACTGTCTGCCAGCTGGCAGCGGGAACGGGACGGG	1351
Qy		
1195	CCAG---AGTTCGAAGTTCCAAAACCAAGTCAAGCCCTCTGAACCTGAGCA-----	1241
Db		
1352	AGTAGCAGCTCTCGGAGCAGCTCGGAGGCGAGTGGATTTCTGTGCTCTCAGAGCCACAGCG	1411
Qy		
1242	-----TGCCCCACGCCCTGGACAGACACCTTCTGCCAGGCGCAAGCTGATGGG	1290
Db		
1412	CTCTACCCCGTGGCAAAATAACAGAAATGCCTCTTGCACACTGCCCTGAATGAGATCAGTAC	1471
Qy		
1291	CTCTATCCCAATCTCTCGGNAACGGTCCAGCTTCTACAGCTGTGCAGGGGGCGGCTGTTC	1350
Db		
1472	CAGCAGAACTGCCAGGCCGGGCTTGCTTTCGACACCAAGCTGTGATTGCTGCAACTGG	1528
Qy		
1351	CAGCAAAAGTGTCCCGACAGGCCCTGGTGTTCAGCAAACTCTGCAAAATGCTGCACCTGG	1407
Db		

RESULT 7

```

US-09-039-198A-1
; Sequence 1, Application US/09039198A
; Patent No. 6200951
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; APPLICANT: Tjoelker, Larry W.
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive/6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/039,198A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 27866/34391
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1636 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..1399
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 65..1399
US-09-039-198A-1

```

Query Match 29.6%; Score 481; DB 4; Length 1636;
Best Local Similarity 60.88; Pred. No. 1.8e-131;
Matches 853; Conservative 0; Mismatches 520; Indels 30; Gaps 3;
QY 126 CAGGTCCTGTCCTTATACTGAATTCGAGCTCGGCTCTGCCTACCAAGCTGACATGCTACT 185

Db 24 CAGGTTTCATGGTCTCTGCTGATGCCATGCGGCTCTGCTGCAAAATGCTGCTACT 83
Qy 186 TCACCAACTGGCCAGTACCGGCGAGGCTGGGGGCTTCTCATGCTGACAAATGACCC 245
Db 84 TCACCAACTGGCCAGTACAGACAGGGGAGGCTGCTTCTGCTCCAGGACTGGACC 143
Qy 246 CCTGCTCTGACCCACTGATCTACGCTTGTGGGAGGCGAGAAACAGAGATACCA 305
Db 144 CCAGGCTTTGACCCACTGATCTACGCTTGTGGATGACCAACACAGCTGAGCA 203
Qy 306 CCATCGAATGAACATGATGATCTCTACAGCTTTCAATGGCTGAAATTAAGACA 365
Db 204 CCATGATGATGAATGACAGAGACTCTCTACAGGAGTTCAATGGCTGAAAGATGATC 263
Qy 366 GCCAGCTGAAACATCTCTGCTGCTGAGCTGGAATCTGGGACTGCCCTTTCACTG 425
Db 264 CCAAGCTGAAGACCTGTTAGCAATGAGAGCTGGAATTTGGGACTCAGAGTTCAAG 323
Qy 426 CCATGGTTTCTACTCTGAGAACCCGAGACTTTTCACTACCTCAGTCAATCAATTCCTGC 485
Db 324 ATATGGTAGCCAGCCGCAACACCGTCAGACCTTTGTCACTGCGCCATCAGGTTTCTGC 383
Qy 486 GCCAGTATGATTTGAGGGCTGACCTTTGACTGGGAGTACCTGCTGCTGGAGGCC 545
Db 384 GCAAAATACAGCTTTGACGGCTTGACCTTGACTGGGAGTACCCAGGAGGCCAGGGAGCC 443
Qy 546 CTCCTCAGGCAAGCATCTCTTCACTGCTGCTGCTGAGGAAATCGTGAAGCTTTTGAGC 605
Db 444 CTGCGGTAGACAGAGGCTTTCAACCTGTTGACAGGACTTGGCCAAATGCTTCCAGC 503
Qy 606 AGGAGGCCAAGCAGTCAACAAAGCCGAGGTGATGCTGCTGCTGCTGCTGCTGCTGCA 665
Db 504 AGGAAGCCAGACCTCAGGGAAGGAAGCGCTTCTTCTGAGTGACGCGGTTCCAGCTGGC 563
Qy 666 TCTCCATATCCAGTCTGGCTATGAGATCCCGCAACTGTACAGTACTGCTGCTGCTGCT 725
Db 564 AGACCTATGATGATGCTGGATGAGGTGAGGAGGCAAAATCGCCAGAACTGGATTTGTCA 623
Qy 726 ATGTCATGACTACGACTCCATGCTGCTGCTGAGGAGGCTACAGTGGAGAGAACAGCCCC 785
Db 624 ACCTTATGGCTAGCACTTCCATGCTCTTGGGAGAGGCTACAGGACATACAGCCCC 683
Qy 786 TCTCAAAATCCGCACTGACACCGGAGAGACCGCTTACCTCAATGTGGATATGATGATGA 845
Db 684 TCTCAAGAGGCAAGAGAGAGTGGTGCAGCAGCGCTCAACGCTGATGCTGCTGCTGCT 743
Qy 846 ACTACTGGAAGCAATGGAGGACAGCTGAGAGCTCATGCTTGGATTCCTTACCTATG 905
Db 744 AACAGTGGCTGCAAGAGGAGGACCCCTGCCAGCAAGCTGATCCTTGGCATGCTTACCTACG 803
Qy 906 GACACAACTTCTCCTGAGCAACCCCTCCAAACACTGGAATGTTGGTCCCGCCACCTTCTGGTG 965
Db 804 GAGCCTCTTCACTGCGCTCTCTCATCAGACACAGAGTGGGGCCCGCCAGCCAGGCT 863
Qy 966 CTGCTCTGCTGGGCCCTATGCAAGAGGCTGCTGGGATCTGGGCTTACTAGAGATCTGTA 1025
Db 864 CTGGCACTCCAGGCGCCCTTCCAAAGGAAGGAGGATGCTGGCTTACTATGATGATGCT 923
Qy 1026 CTTCTCTGAAATGAGGACACTCAGGATGGGATGCCCTCAGGAAGTGCCTTATGCT 1085
Db 924 CCI-----GGAAGGGGGCCCAACAGAGATCCAGGATCAGAGGTCGCCCTTACATCT 977
Qy 1086 ATCAGGCAATGTGTGGTTGGCTTATGACAACTCAAGAGCTTCGATATTAAGGCTCAAT 1145
Db 978 TCCGGGCAACAGGAGTGGTGGCTTGTGATGATGGAGAGCTTCAAAACAGGCTCAGCT 1037
Qy 1146 GGCTTAAGCAACAAATTTGGAGGCCCATGCTGGGCCATGATCTGGATGACTTCA 1205
Db 1038 ATCTGAAGCAGAGAGGAGTGGGGGGCCCATGCTGGGCACTGGACTTAGATGACTTGG 1097
Qy 1206 CTGGCACTTTCTGCAACAGGCAAGTTCCTCCCTTAATCTCCACCTGAAAGGCCCTCG 1265

RESULT 8

US-08-877-599-1

; Sequence 1, Application US/08877599

; Patent No. 6372212

; GENERAL INFORMATION:

; APPLICANT: Gray, Patrick W.

; TITLE OF INVENTION: Chitnase Materials and Methods

; NUMBER OF SEQUENCES: 17

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

; STREET: 6300 Sears Tower, 233 South Wacker Drive

; CITY: Chicago

; STATE: Illinois

; COUNTRY: United States of America

; ZIP: 60606-6402

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/877,599

; FILING DATE:

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/663,618

; FILING DATE: 14-JUN-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Rin-Laures, Li-Hsien

; REGISTRATION NUMBER: 33,547

; REFERENCE/DOCKET NUMBER: 27866/33994

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 312/474-6300

; TELEFAX: 312/474-0448

; TELEX: 25-3856

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1636 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: CDNA

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 2..1399

; FEATURE:

; NAME/KEY: mat_peptide

; LOCATION: 65..1399

; US-08-877-599-1

Query Match 29.6%; Score 481; DB 4: Length 1636;
Best Local Similarity 60.8%; Pred. No. 1.8e-131;
Matches 853; Conservative 0; Mismatches 520; Indels 3

QY	126	CAGGTCCTTGCTCTTACTAGTAATTTGGAGCTCGCGCTCTGCCTACCAAGCTGACATGCTACT	185
Db	24	CAGGTTTCATGGTCTGCTGCTGATGATCCCATGGGGCTCTGCTGCAAAACCTGGTCTGCTACT	83
QY	186	TCACCAACTGGGCCCACTACCGGCCAGGCTTGGGGCGCTTCATGCGCTGCAACAATCGACC	245
Db	84	TCACCAACTGGGCCCACTACAGACAGGGGAGGCTGCTTCTTCCCAAGGACTTGGACC	143
QY	246	CTGCTCTCTGTACCCACCTGATCTAGCGCTTTCTCTGGGAGGCAACAACGAGATCACCA	305
Db	144	CCAGCGCTTTGCACCCACCTCATCTAGCGCTTCGCTGCGATGACCAACCAACGCTGAGCA	203
QY	306	CCATCGAATGGAAGGATGTGACTCTCTACCAAGCTTTCATGCGCTGAAAAATAGAACCA	365
Db	204	CCACTGAGTGGAAATGAGAGACTCTCTACCAAGAGTTCATGGCCCTGAAGAAGATGAATC	263
QY	366	GCCAGCTGAAAACCTCTCTGGGCCATTTGGAGCTTGGAACTTTCGGGACTGCCCCCTTTCAC	425
Db	264	CCAAGCTGAAGACCCCTGTTAGCCATCGGAGCTTGGAAATTTTTCGCAAGTTTCACAG	323
QY	426	CCATGGTTTCTACTCTGTAGAACCGCCAGACTTTTCATCACTCAGTCAATCAAAATTCCTGC	485
Db	324	ATATGTFAGCCAGGGCCACAAACCGTCAGACCTTTGTCAACTCGGCCCATCAGGTTTCTGC	383
QY	486	GCCAGTATGAGTTTGACGGCTGGACTTTTCACCTGGGAGTAGCCCTGGCTCTCTGCTGGAGCC	545
Db	384	GCAAAATACAGCTTTGACGGCTTGACCTTGACTGGGAGTACCACAGNAAGCCAGGGAGCC	443
QY	546	CTCCTCAGGACAAAGCATCTCTTCACCTGCTCTGTGTGAGGAAATTCGCTGGAAGCTTTTGAGC	605
Db	444	CTCCCTAGACAAAGGAGCGCTTCACAAACCTGTGTACAGGACTTTGGCCAAATGCTCTCCAGC	503
QY	606	AGGAGGCCAAGCAGATCAACAAGCCAGGCTGATGCTACTGCTGCAAGTAGCTGCTGGCA	665
Db	504	AGGAGCCCAAGACTCAGGAAGAGACGCCCTTCTTCTGAGTGCAGCGGTTTCCAGCTGGGC	563
QY	666	TCCTCAATATCCAGTCTGGCTATGAGATCCCCCAACTGTCAACAGTACTGGACTACATCC	725
Db	564	AGACCTATGTGGATGCTGGATACGAGGTGCACAAATCGCCAGAACCTTGGATTTGTCA	623
QY	726	ATGTCATGACCTAGGACCTCCATGGCTCTCTGGGAGGGCTACACTGAGAGAACAGCCCCC	785
Db	624	ACCTTATGGCCTACGACTTCCATGGCTTCTTGGGAGAAGGTCAACGGGACATTAACAGCCCC	683
QY	786	TCCTACAAATACCCGACTGCACCCGGCAGCAACGCCCTACCTCAATGTGGATTATGTCTATGA	845
Db	684	TCTACAAGGGCAAGAGAGATGGTTCACACGCCAGCCCTCAACGTGGATGCTGCTGTGC	743
QY	846	ACTACTGGAAGGACAATPGAGCACCAAGCTGAGAACTCATCTGTTGGATTCCTTACCTATG	905
Db	744	AACAGTGGCTTGCAGAAGGGACCCCTGCCAGCAAGCTGATCTCTTGGCATGCCCTACCTAOC	803
QY	906	GACACAACTTCATCTCTGAGCAACCCCTCCAAACACTGGAATTTGTTGCCCCCACCTCTGGTG	965
Db	804	GAGCTCTTTCACACTGGCCCTCTCTATCAGACACAGAGTGGGGGGCCCCAGCCACAGGGT	863
QY	966	CTGCTCTGTGGGCCCTATGCCAAGGAGTCTGGGATCTGGGCTTACTACGAGATCTGTA	1025
Db	864	CTGCACTCCAGCCCTTCCACCAAGGAAGGAGGATGCTGGCCCTACTATGAAGTCTGCT	923
QY	1026	CCTTCTGAAAATGAGCCACATCAGGGATGGGATGCCCTCAGGAAGTGCCCTTATGCTCT	1085
Db	924	CCCT-----GGAAAGGGGGCCACCAACAGAGAAATCCAGGATCAGAGGTTGCCCTACATCT	977
QY	1086	ATCAGGCGCAATGTGGTGTGGCTATGACAAACATCAGAGCTTCGATATTAAGGCTCAAT	1145
Db	978	TCGGGACACCAAGTGGGTGGGCTTTGATGATGTGGAGAGCTTCAAAACCAAGGTCAGCT	1037
QY	1146	GGCTTAAGCAACAATAATTTGGAGGGGCCCATGCTTGGGGCCATGTCTGGGACCTTGATCTGGATGACTTCA	1205

RECEIVED

```

US-09-267-574-1
: Sequence 1, Application US/09267574
: Patent No. 6399571
: GENERAL INFORMATION:
: APPLICANT: Gray, Patrick W.
: APPLICANT: Tjoelker, Larry W.
: TITLE OF INVENTION: CHITINASE CHITIN-BIND
: FILE REFERENCE: 27866/35407
: CURRENT APPLICATION NUMBER: US/09/267,574-1
: CURRENT FILING DATE: 1999-03-12.
: EARLIER APPLICATION NUMBER: 09/039,198
: EARLIER FILING DATE: 1998-03-12
: NUMBER OF SEQ ID NOS: 39
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 1
: LENGTH: 1636
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (2)..(1399)
: FEATURE:
: NAME/KEY: mat_peptide
: LOCATION: (65)..(1399)
: US-09-267-574-1

```

Query Match 29.6%; Score 481; DB 4; Length 1636;

[illegible]

Db 204 CCACTGAGTGAATGACGAGACTCTCTACAGGAGTTCATGGCTGAAGAAGATGATC 263
QY GCCAGCTGAAACACTCTCTGCGCATTTGGAGCTGGAACCTTCGGAGCTGCCCTTTCACTG 425
Db 264 CCAAGCTGAAGACCTGTACCCATCGAGGCTGGAATTCGGCACTCAGAAGTTACAG 323
QY CCATGGTTTCTACTCTGAGAACGCCAGACTTTTCATCACCCTCAGTCAAAATTCCTGC 485
Db 324 ATATGGTAGCCAGCCCAACACCTGACAGCTTTGTCAACTCGGCCATCAGTTTCTGC 383
QY GCCAGTATGAGTTTACGGGCTGACCTTTGACTGGGAGTACCCTGGCTCTCTGTTGGAGCC 545
Db 384 GCAATACAGCTTTGAGCGGCTTACCTTTGAGTGGAGTACCCAGGAGCCAGGGAGCC 443
QY CTCCTCAGCAAGCATCTCTCTACTGCTCTGCTGAGGAAATCGTGAAGCTTTTGAGC 605
Db 444 CTGCGTAGACAAGAGGCTTCAACACCTTGTACAGGACTTGGCCATGCCCTTCAGC 503
QY AGGAGGCCAAGCAGATCAACAGCCAGGCTGTGTCAGTCTGCTGCAAGTCTGCTGCA 665
Db 504 AGAAGCCAGACCTCAGGGAAGGAAGCCCTTCTCTGAGTGCAGCGGTTCCAGCTGGC 563
QY TCTCAATATCAGTCTGGCTATGATCCCTCCACTGTCAGTACCTGACCTGACATCC 725
Db 564 AGACTATGTGGATGCTGGATACGAGGTGGACAATCGCCCAAGCTGGATTTGTCA 623
QY ATGTCATGACCTACGACTCCATGGCTCCTGGAGGCTACACTGGAGAACAGCCGCC 785
Db 624 ACCTATGGCTACGACTTCCATGGCTTGGAGAGGTCACGGGACATACAGCCGCC 683
QY TCTCAATATCAGTCTGACACCGGAGCAAGCCCTTCTCTGAGTGCAGCGGTTCCAGCTGGC 845
Db 684 TCTCAAGAGCAAGAGAGAGTGTGTCAGCAGCAGCCCTCAACGCTGGATGCTCTGCTGC 743
QY ACTAGTGAAGCAATGAGCACCAGCTGAGAGCTCATGTTGGATTCCTACCTATG 905
Db 744 AACAGTGTGCAAGAGGAGCCCTGCCAGCAAGCTGATCTTGGCATGCTACCTACG 803
QY GACAACTTCTATCTGAGCAACCCCTCAACACTGGAATGTGTCGCCCAACCTCTGCTG 965
Db 804 GACGCTCTTCACTGAGCTTCTTATCAGACACAGAGTGGGGGCCGCCACAGGT 863
QY CTGCTCTGCTGGCCCTATGCAAGAGTCTGGGATCTGGGCTTACTAGAGATCTGTA 1025
Db 864 CTGCACTTCCAGGCCCCCTTCAACAAAGAGAGGAGTCTGGCTTACTATGAGTCTGCT 923
QY CTTCTCTGAAATGAGGCACTCAGGATGGGATGCCCTCAGGAAGTGCCTTATGCT 1085
Db 924 CCT-----GGAAGGGGCCACCAACAGAGATCCAGGATCAGAGTGGCCCTACATCT 977
QY ATCAGGCAATGTGTGGTGGCTATGACAACTCAAGAGCTTGGATATTAAGGCTCAAT 1145
Db 978 TCCGGACACCACTGGTGGCTTTCATGATGGAGAGCTTCAAAACCAAGTCACT 1037
QY GGCTTAGCACAACAAATTTGGAGGCCATGCTGGGCAATGATCTGGAGACTTCA 1205
Db 1038 ATCTGAAGCAGAGGACTGGGGGCCATGCTGGGCACTGGACTTAGATGACTTTG 1097
QY CTGCACTTCTGCAACAGGCAAGTTTCCCTTAATCTCCACCTCAAGAAGGCCCTG 1265
Db 1098 CCGCTTCTCTGCAACAGGCGGATACCCCTTATCAGACGCTACGGCAGGAAGTGA 1157
QY GCCTGAGAGTGAAGTTGACCGGCTCCAGCTCAGGCCATTTAGGCCAATAACTGCTCTC 1325
Db 1158 GTCTTCCATATCTTGCCTTACAGGACCCAG---AGCTTGAAGTTCCAAACACAGTTCAGC 1214
QY CCAGTGCAGCGGGAACGGAGCGGGAGTAGCAGCTCTGGAGGAGCTCGGGAGGAGTG 1385
Db 1215 CTTCTGAACCTGAGCA-----TGCCCGGAGCCCTGGAGAGACA 1253
QY GATTCTGTGCTGTCAGAGCAACGGCTCTACCCGCTGGCAATTAACAGAAATGCCCTTCT 1445
Db 1254 CGTTCTGCCAGGCAAGCTGATGGGCTCTATCCCAATCTCTCGGGAACGGTCCAGCTTCT 1313

QY 1446 GGCACTGCGTGAATGAGTACACAGCAGAACTGCAGGCGCGGCTTCTCTTCGACA 1505
Db 1314 ACAGCTGTGAGCGGGGCGCTGTTCAGCAAAAGCTGCCGACAGGCTGGTGTTCAGCA 1373
QY 1506 CCAGCTGTGATGCTGCAACTGG 1528
Db 1374 ACTCTGCAATGCTGCACCTGG 1396

RESULT 10.

US-08-486-839-5
; Sequence 5, Application US/08486839
; Patent No. 5928928
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: A human chitinase, its recombinant
; production, its use for decomposing chitin, its use
; TITLE OF INVENTION: In therapy or prophylaxis against infection diseases.
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11758
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,839
; FILING DATE: 07 - June - 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Baron, Ronald J.
; REGISTRATION NUMBER: 29,281
; REFERENCE/DOCKET NUMBER: 294-26
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1713 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; US-08-486-839-5

Query Match 29.4%; Score 477.2; DB 2; Length 1713;
Best Local Similarity 62.7%; Pred. No. 2.4e-130;
Matches 762; Conservative 0; Mismatches 448; Indels 6; Gaps 1;
QY 92 CTGACTGCAACCATGACAAAGCTTATTCCTCAGAGTCTTCTCTTACTGAAATTTG 151
Db 1 CTGAGCTGCAATGATGTCGGCTCTGTGGCTGGCAGGTTTCATGGTCTGCTGATGATC 60
QY 152 CAGCTCGGCTCTGCCCTACCAGTGCATCTTACCACAACTGGGCCAGTACCGGCCA 211
Db 61 CCATGGGCTCTGCTCCAAAGCTTGTCTACTTCCACAACTGGGCCAGTACAGACAG 120
QY 212 GGCTGGGGGCTTCATGCTGACACATCGACCCCTGCTCTGTACCCACCTGATCTAC 271
Db 121 GGGGAGCTGCTCTCTCCAGAGGACTTGGACCCCGAGCTTTGCACCCACCTCATCTAC 180
QY 272 GCCTTTGCTGGGAGGAGAACACAGATCACCACCATCGAATGGGAACAGATGTGACTCTC 331
Db 181 GCCTTCTGCTGGCATGACCAACACAGCTGAGCAGCAGTGGATGACAGAGACTCTC 240

Db 421 CTTGACTGGAGTACCCAGGAGCCAGGAGGAGCCCTGCCGTAGACAGGAGCGCTTCACA 480
Qy 572 GTCTGTGTGAGGAATCGTGAAGCTTTTTCAGCAGGAGGAGCCAGAGATCAACAAGCCC 631
Db 481 ACCCTGTACAGGACTTGGCCATGCTTCCAGCAGGAGGAGCCAGACCTCAGGAGGAA 540
Qy 632 AGGCTGTGTGACTGCTGTCAGTAGCTGCTGGCATCTCCAAATATCCAGTCTGGCTATGAG 691
Db 541 CGCCTTCTTGTGAGTGCAGGGTTCACAGTGGGAGACCTATGTGGATGTGGATAGAG 600
Qy 692 ATCCCCCAACTGTACAGTAGTACCTGGACTACATCATCATGACCTAGCAGCTCCATGGC 751
Db 601 GTGGACAAATCGCCAGAACCTGGATTTGTCAACCTTATGGCTACGACTTCATGGC 660
Qy 752 TCCTGGAGGGCTACACTGGAGAGAACAGCCCTCTTCAAAATACCCGACTGACACCGGC 811
Db 661 TCTTGGAGAGAGGTTCAGGGGACATACAGCCCTCTTCAAGAGGCAAGAGAGAGTGT 720
Qy 812 AGCAACCCCTACCTCAATGTGGATTATGTATGAATCTACTGGAAGGACAAATGGAGCACCA 871
Db 721 GCAGCAGCCAGCTCAACGTGGATGCTGTGTGCAACAGTGGCTGCAGAGGGGACCCCT 780
Qy 872 GCTGAGAGCTCATCTGTGATTCCTTACCTATGACACAACTTCATCTGAGCAACCCC 931
Db 781 GCCAGCAGCTGATCTTGGCTGCTTACCTACGAGCTCTTCACTAGCTGGCTTCCTCA 840
Qy 932 TCCAAACACTGGAATGTGTGCTCCCTCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 991
Db 841 TCAGACACCAAGAGTGGGGGCCCCAGCCAGGAGTCTGGCACTCCAGGCCCTTTCACCAAG 900
Qy 992 GAGTCTGGGATCTGGCTTACTAGGAGATCTGATCTTCTGTAATAATGGAGGCACTCAG 1051
Db 901 GAAGAGGGATGTGGCTTACTATGATGATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 954
Qy 1052 GGATGGGATCCCTCAGGAAGTGCCTTATGCTCTATCAGGCAATGTGGGTGGCTAT 1111
Db 955 CAGAGATCCAGGATCAGAGGTGGCTTACATCTTCGGGACAAACAGTGGGTGGCTTT 1014
Qy 1112 GACAACATCAGAGCTTCGATATTAAGGCTCAATAGGCTTAAGCACACAAATTTGGAGGC 1171
Db 1015 GATGATGTGAGAGCTTCAAAACCAAGTCAAGTCACTATCTGAAGCAGAGGAGTGGCGGG 1074
Qy 1172 GCCATGCTGGGCTGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1231
Db 1075 GCCATGCTGGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1134
Qy 1232 TTTCCCTTATCTCCACCTGGAAGAGGCGCTCGGCTCGAGAGTGCAGTTCACCGGCT 1291
Db 1135 TACCCCTCATCCAGAGCTACGCGAGGAACTGATGGGTAAAGCCTTAAGCTTAACTGCTG 1194
Qy 1292 CCAGCTCAGCCCATTTG 1307
Db 1195 CATGTGAGGCGAGGTG 1210

RESULT 12

US-09-343-623-5
; Sequence 5, Application US/09343623
; Patent No. 6303118
; GENERAL INFORMATION:

APPLICANT:
TITLE OF INVENTION: A human chitinase, its recombinant
production, its use for decomposing chitin, its use
in therapy or prophylaxis against infection diseases.

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: Hoffmann & Baron
STREET: 350 Jericho Turnpike

CITY: Jericho

STATE: New York

COUNTRY: United States of America

ZIP: 11758

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/343,623
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/486,839
FILING DATE: 07-June-1995
ATTORNEY/AGENT INFORMATION:
NAME: Baron, Ronald J.
REGISTRATION NUMBER: 29,281
REFERENCE/DOCKET NUMBER: 294-26
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 822-3550
TELEFAX: (516) 822-3582
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1713 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: cdna
HYPOTHETICAL: NO
US-09-343-623-5

Query Match 29.4%; Score 477.2; DB 4; Length 1713;

Best Local Similarity 62.7%; Pred. No. 2.4e-130;

Matches 762; Conservative 0; Mismatches 448; Indels 6; Gaps 1;

Qy 92 CTGACTGCAACACCATGACAAAGCTTATTCCTCTCACAGGCTCTGTCTCTTATCTACTGAATTTG 151
Db 1 CTGAGCTGCATCATGTGTGGCTCTGTGGCCCTGGGAGGTTTCATGTGCTGCTGATGATC 60
Qy 152 CAGCTGGCTCTGCTCTACAGCTGACATGCTACTTCCAACTGGGCCCCAGTACCGGCCA 211
Db 61 CCATGGGCTCTGCTCTCAAAACTGGTCTGCTACTTCCAACTGGGCCCCAGTACAGACAG 120
Qy 212 GGCTTGGGGGCTTCTGCTGCTGACAACTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 271
Db 121 GGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
Qy 272 GCCTTGTCTGGGAGGAGAACACAGATCACCACCATCGAATGGAACGATGTGCTCTC 331
Db 181 GCCTTGTCTGGGAGGAGAACACAGATCACCACCATCGAATGGAACGATGTGCTCTC 240
Qy 332 TACCAAGCTTCAATGGCTGAAAATAAGAACAGCAGCTGAAAACCTCTCTGCGCATTT 391
Db 241 TACCAAGGATTCATGGCTGAAAATAAGAACAGCTGAAAACCTCTCTGCGCATTT 300
Qy 392 GGAGGCTGGAACCTCGGAGCTGCCCTTTTCCACTGCGCATGTTCTTCTACTCTGAGAACCGC 451
Db 301 GGAGGCTGGAATTCGGCACTCAGAGTTTCCAGATATGTTAGCCAGGCCCCAACACCGT 360
Qy 452 CAGACTTTCATCACCCTCAGTTCATAAATTCCTGCGCAGTATGAGTTTTCAGCGGCTGGAC 511
Db 361 CAGACCTTCTCAACTCGGCCATCAGTTTCTGCGCAATATACAGCTTTGACGCGCTTGAC 420
Qy 512 TTTGACTGGGAGTACCTGCTGCTGCTGCGGAGCCCTCTCTCAGGACAGCATCTCTTCACT 571
Db 421 CTTGACTGGGAGTACCCAGGAGCCAGGAGGCGCTCTGCTGAGGAGGAGGAGGAGGAGGAG 480
Qy 572 GTCTCTGTCAGGAAATGCTGAAAGCTTTTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 631
Db 481 ACCCTGGTACGAGCTTGGCAATGCTTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
Qy 632 AGGCTGATGCTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 691
Db 541 CGCCTTCTTCTGAGTGCAGCGGTTCCAGCTGGGAGACCTATCTGGATGCTGGATAGGAG 600
Qy 692 ATCCCCCAACTGTACAGACTACCTGGAGTACATCCATGCTATGCTACCTACGACCTCCATGGC 751

Db 601 GTGGACAAAATGCCAGAACCTGGATTTTGTCAACCTTATGSCCTACGACTTCCATGCG 660
QY 752 TCTGGGAGGGCTACACTGGAGAGACAGCCCTCTCAAAATACCAGGACTGACACCGGC 811
Db 661 TCTGGGAGAGGTCAAGGAGATACAGAGCCCTCTCAAGAGGCAAGAGAGTGGT 720
QY 812 AGCAAGCGCTACTCTCAATGGATTATGTCACTGAACTTCTGGAAGGACAAATGAGCACCA 871
Db 721 GCAGAGCCAGCGCTCAACTGATGTGTGTGCAACAGTGGCTGAGAGAGGGAGCCCT 780
QY 872 GCTGAGAAGCTCATCTGTTGGATTCCCTACCTATGGACACAACTTCACTCTGAGCAACCC 931
Db 781 GCAGCAAGCTGATCTTGGCATGCTACTACGGAGCGTCTCTTCACTACCTGGCTCTCTCA 840
QY 932 TCCACACACTGGAATTTGGTGGCCCGCCACCTCTGGTGTGCTGCTGCTGGGCGCTATGCCAAG 991
Db 841 TCAGACACAGAGTGGGGGCCCGCCAGCCAGAGGCTCTGGACACTCCAGGCCCTTCAACCAAG 900
QY 992 GAGTCTGGGATCTGGGCTTACTACGAGATCTGTACCTTCTGAAAAATGGAGCCACTCAG 1051
Db 901 GRAGGAGGATGCTGGCTTACTATGAAGTCTGCTCT-----GGAAGGGGGCCCAAA 954
QY 1052 GATGGGATGCCCTCAGGAAGTGCCTTATGCTATCAGGGCAATGTGGGTGGCTAT 1111
Db 955 CAGAGAATCCAGGATCAGAAGGTGCCCTACATCTTCCGGGACAAACAGTGGGTGGCTT 1014
QY 1112 GACAACATCAAGAGCTTCATATTAAGGCTCAATGGCTTAAGCAACAACAAATTTGAGGC 1171
Db 1015 GATGATGTGGAGAGCTTCAAAACCAAGGTGAGTATCTGAAGCAGAAGGAGTGGGCGGG 1074
QY 1172 GCATGCTGGGCCATGATCTGGATGACTTCACTGGCACTTTCTGCAACACAGGCAAG 1231
Db 1075 GCAATGCTGGCACTGACTTAGATGACTTTGCGGCTTCTCCTGCAACACAGGCGCA 1134
QY 1232 TTTCCCTAATCTCCACCTGAAGAGGCCCTCGGCTCGAGAGTGCAGATTCAGAGCGGT 1291
Db 1135 TACCCCTCATCAGACGCTACGGCAGAACTGAATGGGTAAAGCCTTAACCTGCTGTCA 1194
QY 1292 CCAGCTCAGCCCATG 1307
Db 1195 CATGTGAGGCCAGGTG 1210

RESULT 13
US-08-694-915-1
; Sequence 1, Application US/08694915
; Patent No. 5811535
; GENERAL INFORMATION:
; APPLICANT: Adamou, Julie
; APPLICANT: Kirkpatrick, Robert
; APPLICANT: Rosenberg, Martin
; TITLE OF INVENTION: HUMAN CARTILAGE GP39-LIKE GENE
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-2799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/694,915
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:

ATTORNEY/AGENT INFORMATION:
; NAME: Han, William T
; REGISTRATION NUMBER: 34,344
; REFERENCE/DOCKET NUMBER: ATG50017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5219
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1433 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; US-08-694-915-1

Query Match 22.7%; Score 369; DB 1; Length 1433;
Best Local Similarity 59.8%; Pred. No. 1.4e-98;
Matches 703; Conservative 0; Mismatches 445; Indels 27; Gaps 4;

QY 99 CAACCATGACAAAGCTTATTCTCTCACAGGTCTTCTCTCTTACTGAAATTTTCACGCTCG 158
Db 47 CCACCATGAGCAGAAAGTCTCTCTGGCAGGTGTAGTGGTCTTCTGCTTCTCCAGGGAG 106
QY 159 GCTCTGCCTACCAGCTGACATGCTACTTCCACCAACTGGGCCAGTACCGGCCAGGCGCTG 218
Db 107 GATCTGCCTACAACTGGTTTGTCTACTTTACCAACTGGTCCCAAGGACCGCAGCAACCG 166
QY 219 GGGGCTTCATGCTGACACATCGACCCCTGCTGTACCCACCTGATCTAGCCCTTTG 278
Db 167 GAAATTTACCCCTGAGAATATTGACCCCTTCTATGCTCTCATCTATTCATTCATTCG 226
QY 279 CTGGGAGGAGACAAACGAGATCACCACCATCAATGGAACGATGTGACTCTCTACCAAG 338
Db 227 CCAGCATCGAAACAAACAGGTATCATCAAGGACAGAGTGAAGTGTGCTCTACCA 286
QY 339 CTTTCAATGGCTGAAAAATAAGAACAGCAGCTGAAACTCTCTCTGGCCATTTGAGGCT 398
Db 287 CCATCAACAGTCTCAAAACCAAGAAATCCCAAAATTTCTCTGCTCATTCATTCGAGGCT 346
QY 399 GGNACTTCGGGACTGCCCTTTTCACTGCCATGTTTCTACTCTCTGAGAACCCGACACTT 458
Db 347 ACCTGTTGGTTCCAAAGGGTTCCACCCTATGGTGGATTCTTCTACATCAGCTTGAAT 406
QY 459 TCATCACTCAGTCATCAAAATCTCTGCGCAGTATGAGTTTGACGGCTGGACTTTGACT 518
Db 407 TCATTAACCTCCATATCTCTGTTCTGAGGACCACTAATCTTGTGAGCTGGATGTAAGCT 466
QY 519 GGGAGTACCCTGCTCTCTGTTGGAGCCCTCTCAGGACCAAGCATCTCTTCACTGCTCTG 578
Db 467 GGATCTACCCAGA-----TCAGAAAGAAACAACTCATTTCTACTGTGCTGA 511
QY 579 TGCAGGAAATGCGTGAAGCTTTTGACGAGGAGCCCAAGCAGATCAACACCCAGGCTGA 638
Db 512 TTCATGAGTTAGCAGAGGCTTTTCAGAGGACTTTCACAAATCCCAAGGAAGGCTTC 571
QY 639 TGGTCACTGCTGAGTAGTGTGGCATCTCCAATATCCAGTCTGGCTATGAGATCCCC 698
Db 572 TCTTGAAGTGGGCGGTATCTGAGGAGGCAAAATGATGATACAGCTATCAAGTTGAGA 631
QY 699 AACTGTACAGTACCTGGACTACATGTCATGACCTACGACCTCCATGCTGCTCTGGG 758
Db 632 AACTGGCAAAAGATCTGGATTTCATCAACCTCTCTCTTGTGCTTCTGCTTCTGCTGCTGG 691
QY 759 A-----GGGCTACACTGGAGAGAACACCCCTCTACAAATACCCGACTGACACCGCA 812
Db 692 AAAAGCCCTTATCTACTGGCCACACAGCCCTCTGAGCAAGGGGTGCGAGGAGGCG 751

1052 GGATGGATGCCCTCAGGAAGTCCCTATGCTATCAGGGCAATGTGTGGTGGCTAT 1111
1078 ACCGGCTCCAGATCAGAGGTTCCTACCGAGTCAAGGGGAACAGTGGGTGGCTAT 1137
1112 GACAACATCAAGAGCTTCGATATTAAGGCTCAATGGCTTAAGCAACAATAATTTGGAGGC 1171
1138 GATGATGTGAAGAGTATGGAGACCAAGGTTCAGTTCTTAAAGAATTTAAACCTGGGAGGA 1197
1172 GCATGGTGGGCCATTGATCTGGATGACTTCACTGGCACTTTCTGCAACAGGCAAG 1231
1198 GCATGATCTGGTCTTATGACATGGATGACTTCACTGGCAATCTTCAACAGGCGCCT 1257
1232 TTTCCCTCAATCTCCACCTCGAAGAAGGCCCTCGGC 1267
1258 TACCCTCTGCTCAAGCAGTCAAGAGAGCCTTGGC 1293

RESULT 15
PCT-US94-07754-4
; Sequence 4, Application PC/TUS9407754
; GENERAL INFORMATION:
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
; TITLE OF INVENTION: ASSAY FOR YKL-40 AS A MARKER FOR
; TITLE OF INVENTION: DEGRADATION OF MAMMALIAN CONNECTIVE TISSUE MATRICES
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SPENSLEY HORN JUBAS & LUBITZ
; STREET: 1880 CENTURY PARK EAST, FIFTH FLOOR
; CITY: LOS ANGELES
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/07754
; FILING DATE: 08-JUL-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: HOWELLS, STACY L.
; REGISTRATION NUMBER: 34,842
; REFERENCE/DOCKET NUMBER: FD 3665
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/455-5100
; TELEFAX: 619/455-5110
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1681 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: YKL-40
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 135..1681
PCT-US94-07754-4

Query Match 20.9%; Score 339.8; DB 5; Length 1681;
Best Local Similarity 58.9%; Pred. No. 6e-90;
Matches 669; Conservative 0; Mismatches 442; Indels 24; Gaps 4;

125 ACAGGCTCTGCTCTTACTGAAATTTGCAGCTCGGCTCTGCCCTTACCAGCTGACATGCTAC 184
93 ACAGGCTCTGCTCTGCTGCTGCTCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 152
185 TTCACCACTGGGCCAGTACCGGCCGCTGGGCGCTTTCATGCTGCAACATCGAC 244

Search completed: July 3, 2003, 08:43:18
Job time : 83.4127 secs

Result No.	Score	Query		Length	DB	ID	Description
		Match					
1	1625	100.0	1625	9	US-10-004-219B-2		Sequence 2, Appli
2	1041	64.1	1525	9	US-10-004-219B-3		Sequence 3, Appli
3	708.8	43.6	1038	9	US-10-202-436A-14		Sequence 14, Appl
4	349.4	21.5	1925	9	US-10-097-340-44		Sequence 44, Appl
5	348.4	21.4	1391	10	US-09-822-830A-402		Sequence 402, Appl
6	342	21.0	449	10	US-09-960-352-678		Sequence 678, Appl
7	338.2	20.8	1681	10	US-09-215-077A-4		Sequence 4, Appli
8	338.2	20.8	1681	10	US-09-262-213A-4		Sequence 4, Appli
9	281.2	17.3	1474	10	US-09-755-231A-8		Sequence 8, Appli
10	139.4	8.6	415	10	US-09-960-352-2589		Sequence 2589, Appl
11	136.8	8.4	414	10	US-09-960-352-13050		Sequence 13050, A
12	136	8.4	418	10	US-09-960-352-3072		Sequence 3072, Ap
13	135.2	8.3	410	10	US-09-960-352-7364		Sequence 7364, Ap
14	135.2	8.3	424	10	US-09-960-352-3057		Sequence 3057, Ap
15	122	7.5	411	10	US-09-960-352-14919		Sequence 14919, A
c 16	120.2	7.4	399	10	US-09-960-352-265		Sequence 265, Appl
17	113.6	7.0	427	10	US-09-960-352-3315		Sequence 3315, Appl
c 18	112.4	6.9	384	10	US-09-960-352-5721		Sequence 5721, Appl
19	112	6.9	426	10	US-09-960-352-4979		Sequence 4979, Appl

```
Db 61 AGAACAATAAAGCTCTGCGGACTGGTGTGACTGCAACCAATGACAAAGCTTATTCT 120
Qy 121 CCTCAGAGTCTTGCTTACTGAATTTTCAGCTCGGCTCTGCTACCAAGCTGACATG 180
Db 121 CCTCAGAGTCTTGCTTACTGAATTTTCAGCTCGGCTCTGCTACCAAGCTGACATG 180
Qy 181 CTACTTCACCAACTGGGCCAGTAGTACCGGCCAGGCTGGGGCGCTTCATGCTGACAACAT 240
Db 181 CTACTTCACCAACTGGGCCAGTAGTACCGGCCAGGCTGGGGCGCTTCATGCTGACAACAT 240
Qy 241 CGACCCCTGCTCTGTACCCAGCTGATCTAGCCCTTTGCTGGGAGGAGACACAGAGAT 300
Db 241 CGACCCCTGCTCTGTACCCAGCTGATCTAGCCCTTTGCTGGGAGGAGACACAGAGAT 300
Qy 301 CACCACCATCAATGGGAACGATGTGACTCTCTACCAAGCTTTCAATGSCCTGAAAATAA 360
Db 301 CACCACCATCAATGGGAACGATGTGACTCTCTACCAAGCTTTCAATGSCCTGAAAATAA 360
Qy 361 GAACAGCAGCTGAAACTCTCTGAGAACCCGACACTTTCATCACCCTCAGTCATCAAT 420
Db 361 GAACAGCAGCTGAAACTCTCTGAGAACCCGACACTTTCATCACCCTCAGTCATCAAT 420
Qy 421 CACTGCCATGTTTCTACTCTGAGAACCCGACACTTTCATCACCCTCAGTCATCAAT 480
Db 421 CACTGCCATGTTTCTACTCTGAGAACCCGACACTTTCATCACCCTCAGTCATCAAT 480
Qy 481 CCTGGCCAGTATGAGTTGACGGCTGGACTTTGACTGGGAGTACCTGGCTCTCGTG 540
Db 481 CCTGGCCAGTATGAGTTGACGGCTGGACTTTGACTGGGAGTACCTGGCTCTCGTG 540
Qy 541 GAGCCCTCTCAGACCAAGCAATCTTCACTGTCTGCTGGTGCAGGAATGGGTGAAGCTTT 600
Db 541 GAGCCCTCTCAGACCAAGCAATCTTCACTGTCTGCTGGTGCAGGAATGGGTGAAGCTTT 600
Qy 601 TGACGAGGAGCCAGCAGATCAACAGCCGAGCTGATGCTACTGCTGCTGCTGCTGCTG 660
Db 601 TGACGAGGAGCCAGCAGATCAACAGCCGAGCTGATGCTACTGCTGCTGCTGCTGCTG 660
Qy 661 TGGCATCTCCAATATCAGTCTGCTATGAGTACCCCACTGTCACAGTACCTGGACTA 720
Db 661 TGGCATCTCCAATATCAGTCTGCTATGAGTACCCCACTGTCACAGTACCTGGACTA 720
Qy 721 CATCCATGTCATGACCTACGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 721 CATCCATGTCATGACCTACGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Qy 781 CCCCCTCTACAATATCCGACTGACCGGAGCAACGCTTACCTCAATGCTGATTTATGT 840
Db 781 CCCCCTCTACAATATCCGACTGACCGGAGCAACGCTTACCTCAATGCTGATTTATGT 840
Qy 841 CATGAATCTGGAAGGACCAATGGAGCACCAGCTGAGAAGCTCATCTGCTGCTGCTGCT 900
Db 841 CATGAATCTGGAAGGACCAATGGAGCACCAGCTGAGAAGCTCATCTGCTGCTGCTGCT 900
Qy 901 CTATGGACACAATCTGATCTGACCAACCCCTGCAACACTGGAATTTGGTCCGCCCTC 960
Db 901 CTATGGACACAATCTGATCTGACCAACCCCTGCAACACTGGAATTTGGTCCGCCCTC 960
Qy 961 TGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
Db 961 TGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
Qy 1021 CTGTACCTCTCTGAAAATGGAGCCACTCAGGATGGATGCTGCTGCTGCTGCTGCTGCT 1080
Db 1021 CTGTACCTCTCTGAAAATGGAGCCACTCAGGATGGATGCTGCTGCTGCTGCTGCTGCT 1080
Qy 1081 TGGCTATCAGGCAATGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1140
Db 1081 TGGCTATCAGGCAATGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1140
Qy 1141 TCAATGGCTTAAACACAAATTTGGAGGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
Db 1141 TCAATGGCTTAAACACAAATTTGGAGGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
```

```
Qy 1201 CTTCACTGGCACTTTTGTCAACACAGGCAAGTTTCCCTTAATCTCCACCTGAAAGGC 1260
Db 1201 CTTCACTGGCACTTTTGTCAACACAGGCAAGTTTCCCTTAATCTCCACCTGAAAGGC 1260
Qy 1261 CCTCGGCTCGAGAGTGCAAGTTGCACGGCTCCAGCTCAGCCCATTTAGCCCAATAACTGC 1320
Db 1261 CCTCGGCTCGAGAGTGCAAGTTGCACGGCTCCAGCTCAGCCCATTTAGCCCAATAACTGC 1320
Qy 1321 TGGTCCCAAGTGGCAGCGGAGGAGGAGGAGTAGCAGCTCTCGAGGAGCTCGGAGG 1380
Db 1321 TGGTCCCAAGTGGCAGCGGAGGAGGAGGAGTAGCAGCTCTCGAGGAGCTCGGAGG 1380
Qy 1381 CAGTGGATTTCTGCTGTGACAGCAACGGGCTCTACCCCTGGCAATAACAGAAATGC 1440
Db 1381 CAGTGGATTTCTGCTGTGACAGCAACGGGCTCTACCCCTGGCAATAACAGAAATGC 1440
Qy 1441 CTTCTGCACTGCTGATGAGTACGTACCAAGCAACTGCGAGCGGCTGCTGCTT 1500
Db 1441 CTTCTGCACTGCTGATGAGTACGTACCAAGCAACTGCGAGCGGCTGCTGCTT 1500
Qy 1501 CGACACCACTGCTGATGCTGCAACTGGGCAATAACCTGACCTGCTATATTCCTCTAG 1560
Db 1501 CGACACCACTGCTGATGCTGCAACTGGGCAATAACCTGACCTGCTATATTCCTCTAG 1560
Qy 1561 GTTCCAGTCTCTTTGCTTAGGACATGTTGCCCTAGCTAAAGTCTGCAATAAAATCAG 1620
Db 1561 GTTCCAGTCTCTTTGCTTAGGACATGTTGCCCTAGCTAAAGTCTGCAATAAAATCAG 1620
Qy 1621 CAGTC 1625
Db 1621 CAGTC 1625
```

RESULT 2

```
US-10-004-219B-3
; Sequence 3, Application US/10004219B
; Publication No. US20030087414A1
; GENERAL INFORMATION:
; APPLICANT: Macrozyme
; APPLICANT: Aerts, Johannes M.F.G.
; APPLICANT: Boot, Rolf G.
; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and
; TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in
; TITLE OF INVENTION: which mucus is involved or infection diseases
; FILE REFERENCE: 2183-5136US
; CURRENT APPLICATION NUMBER: US/10/004,219B
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1525
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Combined DNA/RNA Molecule: mouse
; OTHER INFORMATION: AMCase cDNA sequence and deduced amino acid
; OTHER INFORMATION: sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: mouse AMCase
; OTHER INFORMATION: cDNA sequence and deduced amino acid sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1419)
US-10-004-219B-3
```

```
Query Match 64.1%; Score 1041; DB 9; Length 1525;
Best local similarity 83.4%; Pred. No. 0;
Matches 1198; Conservative 0; Mismatches 230; Indels 9; Gaps 1;
Qy 104 ATCAACAAGCTTATTCTCCCTCAGAGTCTGTCTTATATCTGCTGCTGCTGCTGCT 163
Db 1 ATGCCCAAGCTACTTCTCGTCAAGGCTCGGCTCTTCTGCTGAATGCTCAGCTGGGCTCT 60
```


Qy 1129 CGATATTAGGCTCAATGGCTTAAGCACAACAAATTTGGAGCGCCCTGTCTGGGCCAT 1188
Db 361 CGATATTAGGCTCAATGGCTTAAGCACAACAAATTTGGAGCGCCCTGTCTGGGCCAT 420
Qy 1189 TGATCTGGATGACTTCTACTGGCATTCTGCAACCCAGGGCAAGTTTCCCTTAATCTCCAC 1248
Db 421 TGATCTGGATGACTTCTACTGGCATTCTGCAACCCAGGGCAAGTTTCCCTTAATCTCCAC 480
Qy 1249 CCTGAAGAAGGCCCTCGGGCTGCAGAGTGAAGTTGCACGGCTCCAGCTCAGGCCATTGA 1308
Db 481 CCTGAAGAAGGCCCTCGGGCTGCAGAGTGAAGTTGCACGGCTCCAGCTCAGGCCATTGA 540
Qy 1309 GCCAATAACTGCTGCTCCCAAGTGGCAGCGGGAACGGGAGGAGTACAGCTCTGGAGG 1368
Db 541 GCCAATAACTGCTGCTCCCAAGTGGCAGCGGGAACGGGAGGAGTACAGCTCTGGAGG 600
Qy 1369 CAGCTCGGAGGAGTGGATTCT--GTGCTGTGACAGGCCAACGGCTCTACCCCGTGGCA 1426
Db 601 CAGCTCGGAGGAGTGGATTCTGTGCTTGGCAGAGCAACAGGCTCTAACCCCGTGGG 660
Qy 1427 AATA-----ACAGAAATGCTTCTGGCACTGCGTGAATGGAGTCAACGTACCAG--CAGAA 1479
Db 661 CAATACCAGAGATGCTTCTGGGCACTGCGTGAATGGAGTCAACGTACCAG--CAGAA 720
Qy 1480 CTGCCAGCGGGCTTCT--CTTCGACACAGCTGTGATGCTGCAACTGGGCATPAAACCT 1538
Db 721 TTGCCAGCGGGCTTCTTCTCTGAGACCACTGTGAATGCTGCAACTGGGCATPAAACCT 780
Qy 1539 GACCTGCTATATTCCTCCTAGATCCAGTCTCTTTTGGCTTAGGACATGTTGCC 1594
Db 781 GACCTGCTATATTCCTCCTAGATCCAGTCTCTTTGGCTTAGGACATGTTGCC 836

RESULT 4

US-10-097-340-44
; Sequence 44: Application US/10097340
; Publication No. US20030087250A1
; GENERAL INFORMATION:
; APPLICANT: John MONAHAN
; APPLICANT: Manjula GANNAVAPU
; APPLICANT: Sebastian HOERSCH
; APPLICANT: Shubhangi KAMATKAR
; APPLICANT: Steve G. KOVATS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN
; APPLICANT: Peter VEIBY
; APPLICANT: Gordon B. MILLS
; APPLICANT: Robert C. BAST, Jr.
; APPLICANT: Karen LU
; APPLICANT: Rosemarie SCHMANDT
; APPLICANT: Xumei ZHAO
; APPLICANT: Karen GLATT
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; TITLE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer
; FILE REFERENCE: MRI-030
; CURRENT APPLICATION NUMBER: US/10/097,340
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19

; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 1925
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-097-340-44

Query Match 21.5%; Score 349.4; DB 9; Length 1925;
Best Local Similarity 59.5%; Pred. No. 3.7e-105;
Matches 675; Conservative 0; Mismatches 436; Indels 24; Gaps 4;

Qy 125 ACAGGTCTTGTCTTACTACTGAATTTGCACTCGGCTCTGCTACACAGCTGACATGCTAC 184
Db 148 ACAGGTCTTGTCTTACTACTGAATTTGCACTCGGCTCTGCTACATCAAACTGGTCTGCTAC 207
Qy 185 TTACCAACCTGGGCGGAGTACCGGCGGAGGCTGGGCGCTTCATGCTGACACATGAC 244
Db 208 TACACCACTGGTGGTCCAGTACCGGGAAGGAGTGGGAGCTGCTTCCAGATGGCCCTTGAC 267
Qy 245 CCTGCTCTGTACCCACCTGATCTAGCGCTTCTGCTGGGAGGCAACACGAGATCAC 304
Db 268 CGCTTCTCTGTACCCACATCTACAGCTTGGCCAAATATAGCAACGATCACATGAC 327
Qy 305 ACCATCGAATGGAACAGATGTACTCTCTACCAAGCTTTCAATGGCCTGAAATAAAGAAC 364
Db 328 ACCTGGAGTGGAAATGTGACGCTCTACGGCATGCTCAACACACTCAAGAACAGAAC 387
Qy 365 AGCAGCTGAAACTCTCTGCGCATTTGGAGGCTGGAACTTCGGAGCTGCCCCCTTCACT 424
Db 388 CCCAACCTGAAGACTCTCTTGTCTGCTGGAGGATGGAACTTTGGGTCTCAAGATTTTCC 447
Qy 425 GCCATGTTTCTACTCTGAGAACCGGACAGCTTTTCACTCACTCAGTCATCAATTCCTG 484
Db 448 AAGTAGCTCCCAACACCCAGAGTCCGCGGACTTTCATCAAGTCAGTACCCCATTCCTG 507
Qy 485 CGCAGTATGAGTTTACGGGCTGGAGCTTTGACTGGGAGTACCTGCTCTGCTGGGAGC 544
Db 508 CGCACCATGGCTTTGATGGGCTGGACCTTTCCTGCTTACCTTGGACGAGCA----- 561
Qy 545 CCTCTCAGGACAGCATCTTTCACCTGCTGCTGGTGGAGAAATCGCTGAAGCTTTTGA 604
Db 562 -----GACAAACAGCATTTTACCACCTTAATCAAGAAATGAAGGCCGAATTTATA 612
Qy 605 CAGGAGCCCAAGCAGATCAACAAAGCCAGGCTGATGCTCACTGCTGCAAGTACGTGGC 664
Db 613 AAGGAAGCCAGCAGGAGGAAAAAGC---AGCTCTGCTCAGCGGAGCAGTCTGCTGGGG 669
Qy 665 ATCTCCAATATCCAGTCTGGCTATGAGATCCCCCACTGTCACAGTACCTGGACTACATC 724
Db 670 AAGGTCAACATTGACAGCAGCTATGACATTTGCCAAGATATCCCAACACTGGGATTTTCA 729
Qy 725 CATGTATGACCTACGACCTCCATGGCTCTCTGGGAGGCTTACACTGGAGAGACAGCCCC 784
Db 730 AGCATATGACCTACGATTTTCATGGAGCTGGGCTGGGACCAAGCCATCAGATGCC 789
Qy 785 CTCTACAAATACCGACTGACACCGGCAAGCCCTACCTCAATGTGGATTTATGTGATG 844
Db 790 CTGTTCGAGGTGAGGAGTGAAGTCCCTGACAGATTCAGAACACTGACTATGCTGTG 849
Qy 845 AACTACTGGAAGCAATGGAGCAGCAGCTGAGAAGCTCATGTTGGATTTCCCTACCTAT 904
Db 850 GGGTACATGTTGAGGCTGGGGCTCTCTGCCAGTAAGCTGGTGTGGCATCCCCACCTTC 909
Qy 905 GGACAACTTCACTCTGAGCAACCCCTCCCAACACTGGAATTTGGTGGCCCCCCTCTGCT 964
Db 910 GGGAGGAGCTTCACTCTG---GCTTCTTGAGACTGGTGTGGAGCCCCCAATCTCAGGA 966
Qy 965 GCTGGTCTGCTGGGCCCTATGCCAAGGAGTCTGGGATCTGGGCTTACTACGAGATCTGT 1024
Db 967 CCGGAATTCAGGCGGTTTCCACCAAGGAGGAGGAGCCCTTGCCTACTATGAGATCTGT 1026
Qy 1025 ACCTTCTGAAAAATGGAGCCACTCAGGAGTGGATGGCTGCCCTCAGGAAGTGCCTTATGCC 1084

Db 1027 GACTTCTCT--CGCGGAGCCACAGTCCATAGAACCTCGGCCAGGAGTCCCTATGCC 1083
QY 1085 TATCAGGGAATGTGGGTGGCTATGACAAACATCAAGAGCTTCGATATTAAGGCTCA 1144
Db 1084 ACCAAGGGCAACAGTGGGTAGATACGACGACCAAGAGCGTCAAAAGCAAGGTGCAG 1143
QY 1145 TGGCTTTAAGCACAAACAAATTTGGAGCGGCATGCTGGGCGCAATGATCTGGATGACTTC 1204
Db 1144 TACCTAAGGATAGGAGCTGGCAGCGGCATGATGGCCCTGGACCTGGATGACTTC 1203
QY 1205 ACTGGCAGCTTTTCGAACAGGCAAGTTTCCCTTAATCTCCACCCTGAAGAAGG 1259
Db 1204 CAGGGCTCTTCTGCGCGCAGGATCTGCGCTTCCCTCCCTACCAATGCCATCAAGG 1258

RESULT 5

US-09-822-830A-402
; Sequence 402, Application US/09822830A
; Patent No. US20020142952A1
; GENERAL INFORMATION:
; APPLICANT: Genetics Institute, Inc.
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fechtel, Kim
; APPLICANT: Agostino, Michael J.
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulukota, Kamalakar
; APPLICANT: Graham, James R.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 6402
; CURRENT APPLICATION NUMBER: US/09/822,830A
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/195,604
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 631
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 402
; LENGTH: 1391
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; LOCATION: 1353
; OTHER INFORMATION: n=a,c,g, or t
US-09-822-830A-402

Query Match 21.4%; Score 348.4; DB 10; Length 1391;
Best Local Similarity 59.2%; Pred. No. 6.6e-105;
Matches 680; Conservative 0; Mismatches 441; Indels 27; Gaps 4;
QY 126 CAGGTCTGCTCTATACATGAATTTGCGAGCTCGGCTCGCTACCAGCTGACATGCTACT 185
Db 35 CAGGTGATGGTCTTCTGCTGCTTCTCCAGGAGGATCTGCTCAAACTGGTTGGGACT 94
QY 186 TCACCAACTGGGCGGAGTACCGCCAGGCTGGGCGCTTCATGCTGCAACATCGACC 245
Db 95 TTACCAACTGGTCCCGAGCCGCGCAGGACCAAGAAATTCACCCCTGAGATTTGACC 154
QY 246 CTGCTCTGTATCCCACTGATCTAGCCCTTTCTGGGAGGCGAGAACAGGATCACCA 305
Db 155 CTTCTCTATGCTCTCATCTCATCTATTTCATTCGCCAGCATCGAAACAAACAGGTTATCA 214
QY 306 CCATCGAATGGACGATGTGACTCTCTACCAAGCTTTCAATGGCTGAAATAAGAAC 365
Db 215 TCAAGGACAGAGTGAAGTGAATGCTCTTACCAAGCATTACAGCTCTCAAAACCAAGATC 274
QY 366 GCAGCTGAAAACCTCTCTGGCCATTGGAGCTTGGAACTTTCGGGACTGCCCTTTCACTG 425
Db 275 CCAACTGAAATTCCTTGTCTTCCATTTGGAGGGTACCTGTTTGTTCCTCAAGGGTCCACC 334
QY 426 CCATGTTTCTACTCTCTGAGAACCGCCAGACTTTTCATCAGCTCAGTCAATTCCTGCTC 485

Db 335 CTATGTGGATTTCTTCTACATCAGCTTGGATTTCAATTAATCTCTGTTCTGA 394
QY 486 GCAGATGATGATTTGAGGGCTGAGCTTTTACCTTGGAGTACCTGGCTCTCTCGTGGAGCC 545
Db 395 GGAACCATAACTTTTGTGGACTGGATGTAAGCTGGATCTACCCAG----- 439
QY 546 CTCTCAGGACAAAGCATCTCTTACATCTCTCTGTCGAGGAAATCCGCTGAAGCTTTTGAGC 605
Db 440 ATCAGAAAGAAACACTCATTTTCACTCTGCTGATTCATGATTTAGCAGAAAGCTTTTCAGA 499
QY 606 AGGAGCGCAAGCAGATCAACAAGCCAGGCTGATGTCCTGCTGCTGCTGCTGCTGCTGCTGCTG 665
Db 500 AGGACTTCAAAAATCCACCAAGGAAAGGCTTCTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTG 559
QY 666 TCTCCAATATCCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 725
Db 560 GGCAAATGATTTGATTAACAGCTATCAAGTTGAGAACTGGCAAAAGATCTGGATTTTATCA 619
QY 726 ATGTCATGACTACGACCTCCATGGCTCTCTGGGA-----GGGCTACACTGGAGAGAAC 779
Db 620 ACCTCTGCTCTTGTGACTTCCATGGTCTTTGGGAAAGCCCTTATCACTGGCCACAACA 679
QY 780 GCCCCTCTACAAATACCCGACTGACACCCGAGCAGCAAGCCTACCTCAATGTGGATTTATG 839
Db 680 GCCCTCTGAGCAAGGGGTGGCAGGACAGAGGGCCAAAGCTCTACTACAATGTGGAATATG 739
QY 840 TCATGAACCTGGAAGCAATGGAGCACCAGCTGAGAGCTCATCTGTTGGATTTCCCTA 899
Db 740 CTGTGGGTACTGGATACATGAAGGAATGCCATCAGAGAAGTGGTTCATGGGCATCCCA 799
QY 900 CCTATGACACAACTTTCATCTGAGCAACCCCTCCAACTGGAATTTGGTCCCTCCCT 959
Db 800 CATATGGCACTCTTTCACACTGGCTCTCCAGAAACCCG---TGCGGCCCCCTGCT 856
QY 960 CTGGTGTGCTCTGCTGGGCTCTATGCCAAGAGTCTGGGATCTGGGCTTACTACGAGA 1019
Db 857 CTGGCTCTGGAGCTGCTGGACCATCACAGAGTCTTCAGGCTTCTCGGCTTATATGAGA 916
QY 1020 TCTGTACTTCTCTGAAAATGGAGCCACTCAGGATGGGATGCCCTCAGGAAGTGGCTT 1079
Db 917 TCTGCCAGTTTCTTGA---GGAGCCCAAGATCAGAGGCTCCAGAGTTCAGAGTTCCCT 973
QY 1080 ATGCTATCAGGCAATGTGTGGTGTGGCTTATGACAACTCAAGAGCTTCGATATTAAGG 1139
Db 974 ACCGAGTCAGGGGACCCGCTGGTGGCTTATGATGATGTGAAGATTTGGGCCCCAAG 1033
QY 1140 CTCAATGGCTTAAGCACAACAAATTTGGAGGCGCCATGGTCTGGGCTTACTGATG 1199
Db 1034 TTCAGTTCTTAAAGAAATTTAAAGCTGGGGGTGCTTGTGTTGCTCTTTTGCATGTTG 1093
QY 1200 ACTTCACTGGCACTTCTTCAACCCAGGCAAGTTTCCCTTAATCTCCACCTGGAAGAGG 1259
Db 1094 ACTTCACTGGCAATCTTCAACCCGGGGCTTCCCTCTTGTGTCACAGAGTCAAGAGAA 1153
QY 1260 CCCTCGGC 1267
Db 1154 GCCTTGGC 1161

RESULT 6

US-09-960-352-678
; Sequence 678, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathalaan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AN
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352

QY 1145 TGGCTTAAGCACACAAATTTGGAGCGCCATGCTGGGCCATTTGATCTGGATGACTTC 1204
Db 1089 TACCTGAAGATAGGAGCTGGCAGCGCCATGGTATGGCCCTGGACCTGGATGACTTC 1148
QY 1205 ACTGGGCACTTTGCAACAGGCAAGTTTCCCTTAATCTCCACCTGAAGAAGG 1259
Db 1149 CAGGGCTCTTCTGCGCGCAGGATCTGGCTTCCCTCTCACCATGCCATCAAGG 1203

RESULT 8
US-09-262-213A-4
; Sequence 4, Application US/09262213A
; Patent No. US20020090658A1
; GENERAL INFORMATION:
; APPLICANT: PRICE, PAUL
; APPLICANT: JOHANSEN, JULIA
; TITLE OF INVENTION: ASSAY FOR YKL-40 AS A MARKER FOR DEGRADATION OF
; TITLE OF INVENTION: MAMMALIAN CONNECTIVE TISSUE MATRICES
; FILE REFERENCE: 407T-895412US
; CURRENT APPLICATION NUMBER: US/09/262,213A
; CURRENT FILING DATE: 1999-03-04
; PRIOR APPLICATION NUMBER: US 08/089,989
; PRIOR FILING DATE: 1993-07-09
; PRIOR APPLICATION NUMBER: PCT/US94/07754
; PRIOR FILING DATE: 1994-07-08
; PRIOR APPLICATION NUMBER: US 08/581,527
; PRIOR FILING DATE: 1996-04-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 1681
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-262-213A-4

Query Match 20.8%; Score 338.2; DB 10; Length 1681;
Best Local Similarity 58.9%; Pred. No. 1.8e-101;
Matches 668; Conservative 0; Mismatches 443; Indels 24; Gaps 4;
QY 125 ACAGGCTTCTCTTATCTGAAATTTGCAGCTGGCTCTGCCCTACCAGCTGACATGCTAC 184
Db 93 ACAGGCTTCTCTTATCTGAAATTTGCAGCTGGCTCTGCCCTACCAGCTGACATGCTAC 152
QY 185 TTCACCAACTGGGCCAGTACCGCCAGCGCTGGGCGCTTCATGCTGCAACACATCGAC 244
Db 153 TACACCACTGGTCCAGTACCGGAGCGATGGAGCTGCTCCAGATGCCCTTGAC 212
QY 245 CCCTGCTCTGTACCACTGATCTACGCTTTGCTGGGAGGAGGAGCAACACGAGATCAC 304
Db 213 CGCTTCTGTGTACCACTGATCTACGCTTTGCTGGGAGGAGGAGCAACACGATCAC 272
QY 305 ACCATCGAATGGAACGATGACTCTCTACCAAGCTTCAATGGCTGAAAATAGAAC 364
Db 273 ACCTGGAGTGAATGATGAGCTCTACCGGATGCTCAACACACATCAACACAGAAC 332
QY 365 AGCCAGCTGAAACTCTCTCGGCAATGGAGCTGGAACCTCGGAGCTGCCCTTTCAC 424
Db 333 CCCAACCTGAAGACTCTCTGTGCTGGGAGGATGGAACCTTGGCTCAAGATTTTC 392
QY 425 GCCATGTTTCTACTCTGAGAACCGCCAGCTTTTCATCACTCACTGATCAATTTCTG 484
Db 393 AAGATAGCTTCAACACCCAGAGTGGCGGACTTTTCATCAAGTCAGTACCGCCATTTCTG 452
QY 485 CGCCAGTATGATGAGCGGCTGGACTTTGACTGGGAGTACCTGCTCTGCTGGGAGC 544
Db 453 CGCACCATGCTTGTATGGCGGTGACCTTGCCTGCTTACCTGGACGGAGA----- 506
QY 545 CCTCCTCAGCACAGATCTCTCACTGCTGCTGAGGAAATGCGTGAAGCTTTGAG 604
Db 507 -----CACAAACACATTTTACCACCTTAATCAAGGAATGAAGGCCGAATTTATA 557
QY 605 CAGGAGGCCAAGCAGATCAACAAAGCCAGCGCTGATGGTCACTGCTGAGTAGCTGTGGC 664
Db 1149 CAGGGCTCTTCTGCGGCCAGGATCTGCGCTTCCCTCTCACAATGCCATCAAGG 1203

Db 558 AAGGAAGCCAGCCAGGGAAGAAAGC---AGCTCTGCTCAGCGAGCACTGTCTGCGGGG 614
QY 665 ATCTCAATATTCAGTCTGCTATGAGATCCCACTGTCACAGTACCTGGACTACATC 724
Db 615 AAGTCAACCATTCAGCAGCAGTATGACATTCGCAAGATATCCCAACACCTGGATTTCA 674
QY 725 CATGTCATGACCTACGACCTCCATGCTCTGGAGGGCTTACACTGGAGAGAACAGCCCC 784
Db 675 AGCATCATGACCTACGATTTTCATGGCGCTGGGCGGACACAGCCATCACAGTCCC 734
QY 785 CTCTACAAATPACCCGACTGACACGGCAGCAAGCCCTACCTCAATCTGGATTTATGTCATG 844
Db 735 CTCAGCGAGGCTCAGGAGGATGCAAGTCTCAGAGATTCAGCAACACTGACTATGCTGTG 794
QY 845 AACTACTGGAAGACAAATGAGCAGCAGCTGAGAAGCTCATCGTTGATTCCTTACCTAT 904
Db 795 GGTATCATGTTGAGCTGGGGCTCTCGCAGCTAAGCTGGTGTATGGCATCCCACTTC 854
QY 905 GGACACAACTTCATCTGAGCAACCCCTCCAACACTGGAATTTGGTCCCCCACTCTGCT 964
Db 855 GGGAGGAGCTTCACTCTG---GCTTCTCTGAGACTGGTCTCCAGCGCCAAATCTCAGGA 911
QY 965 GCTGCTCTGCTGGGCCCTATGCAAGGAGTCTGGGATCTGGGCTTACTAGGATCTGT 1024
Db 912 CCGGGAATTCAGGCGCGTTTCAACAAGGAGGAGGACCCCTTGCCTACTATGAGATCTGT 971
QY 1025 ACCTTCTGAAAAAATGGAGCAGCTCAGGGATGGATGCCCTCAGGAAGTGCCTTATGCC 1084
Db 972 GACTTCT---CGCGGAGGCCACAGTCCATAGAACCTCGGCCAGAGTCCCTATGCC 1028
QY 1085 TATCAGGGCAATGTGGTGGCTATGACAACATCAAGAGCTTCGATATTAAGGCTCAA 1144
Db 1029 ACCAAGGGCAACAGTGGTGGTACGACGACAGCAAGGCGTCAAAAGCAAGGTGCAG 1088
QY 1145 TGCGTTAAGCACAAATTTGGAGCGCCATGCTGGGCCATTTGATCTGGATGACTTC 1204
Db 1089 TACCTGAAGATAGGAGCTGGCAGCGCCATGATGGCCCTGGACCTGGATGACTTC 1148
QY 1205 ACTGGCACTTCTGCAACAGGCGCAAGTTCCTCCCTTAATCTCCACCTGAAAGG 1259
Db 1149 CAGGGCTCTTCTGCGGCCAGGATCTGCGCTTCCCTCTCACAATGCCATCAAGG 1203

RESULT 9
US-09-765-231A-8
; Sequence 8, Application US/09765231A
; Patent No. US20020119452A1
; GENERAL INFORMATION:
; APPLICANT: Searle/Monsanto
; APPLICANT: Phippard, Deborah
; APPLICANT: Vasanthakamur, Geetha
; APPLICANT: Dotson, Stanton
; APPLICANT: Ma, Xiao-Jun
; TITLE OF INVENTION: Osteoarthritis tissue-derived nucleic acids, polypeptides,
; TITLE OF INVENTION: vectors, and cells
; FILE REFERENCE: SO-3221 PR
; CURRENT APPLICATION NUMBER: US/09/765,231A
; CURRENT FILING DATE: 2001-01-18
; NUMBER OF SEQ ID NOS: 82
; SEQ ID NO 8
; LENGTH: 1474
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-765-231A-8

Query Match 17.3%; Score 281.2; DB 10; Length 1474;
Best Local Similarity 59.4%; Pred. No. 1.6e-82;
Matches 667; Conservative 0; Mismatches 423; Indels 32; Gaps 10;
QY 158 GGCTCTGCTTACCAGCTGACATGCTACTTACCAGCTGGG-CCCACTACCGCCGAGGCT 216
Db 128 GGATCTGCTTACAACTGGTTGCTACTTTTACCAACTGGGTCCCGAGGACCGGCAAGC 187

Qy	217	GGGGCGCTTCATGCTTGAC - AACATCGAGCCCTCGCTGTGTACCCACCTGATCTACGGCT	275
Db	188	AGAAAAATTCACCCCTCGAGGAATATTGACCCCTTCCTATGCTCTCATCTCATATTTCAT	247
Qy	276	T - TCGTGGGAGGCAGAACACGAGATCACCACCATCGAATCGAAGTGAAGATGCTCACTCTCTAC	334
Db	248	TGCGCCAGCATCGAAAACAACAGGTTATCATCAAGGACAAGATGAAGTGTGCTCTCTAC	307
Qy	335	CAAGCTTTCAA - TGCGCTGAAAAATAAGAACAGCAGCTGAAAACTCTCTCGGCCATTGG	393
Db	308	CAGACCATCAACAGCTTCTCAAAACCAAGAATCCCAAACCTGAAAAATCTCTTGTCCATTGG	367
Qy	394	AGGCTGAACACTCGGGACTGCCCTTTTCACGTGCCATGGTTTCCTACTCTGTAGAACCGCCA	453
Db	368	AGGCTACTGTGTTGGTTCCAAAGGGTTCCACCCCTATGGTGTGATTTCTCTACATCACGGCT	427
Qy	454	GACPTTTCATCACCTTCAGTCATCAAAATCTCGCGCCAGTATGAGTTTGAAGGGCTGGACTT	513
Db	428	GGAAATTCATTAACTTCATATCTCTGTTCTGAGGAACCATAACTTTGATGAGCTGGATGT	487
Qy	514	TGACTGGGAGTACCCTTGCTCTCTGCTGGAGCCCTCTCAGGACAAGCATCTCTTCACTGT	573
Db	488	AAGCTGATCTACCCAG - -----ATCAGAAAAGAAAACACTCATTTCACTGT	532
Qy	574	CCTGGTGCAGGAATCGTGAAGCTTTTGACGAGGGCCCAAGCAGATCAACAGCCCAAG	633
Db	533	GCTGATTCATGAGTTACGAGNAGCCTTTCAGAGAGCTTCAAAAATCCACCAAGGAAG	592
Qy	634	GCTGATGGTCACCTGCTGCA - GTAGCTGCTGGCATCTCCAATATCCAGTCTGGCTATGAGA	692
Db	593	GCTTCTCTTGACTCGGGGGTATCTCGAGGAGGCAATGATTTGATAACAGCTATCAAG	652
Qy	693	TCCCCAACTGTCACGACTGTCGACTACATCCATGTATGATGACCTAGACCTCCATGGCT	752
Db	653	TTGAAAACATGGCAAAAGATCTGGATTTCACTCAACCTCCTGCTCTTGACTTCCATGGGT	712
Qy	753	CCTGGGA - -----GGGCTACACTGGAGAGAACAGCCC - CCTCTACAANAATCCCGACTGAC	805
Db	713	CTTGGGAAAAGCCCTTTATCACCTGGCCACAAACAGCCCTGCTGAGCAAGGGGTGCGAGAC	772
Qy	806	ACGGCAGCAACGCCCTACCTCAATGCTGGATTTATGTCATGAACTACTGGAAGGACAATGGA	865
Db	773	AGAGGGCCAAAGCTCTACTCAATGTGGAAATATGCTGTGGGGTATCTGGATACATAAGGA	832
Qy	866	GCACCACTGAGAAAGCTCATCGTTGGATTCCTTACCTATGAGACACAACCTCATCCTGAGC	925
Db	833	ATGCCATCAGAGAAGTGGTCATGGGGATCCCCACATATGG - GGCATCTTCAACATG	890
Qy	926	AACCCCTCCAACTTGAATTTGGTGGCCCCACCTCTGCTGGTGTGCTGCTGGCCCCCTAT	985
Db	891	GCCTCTCGAAAACCAACCTGGGGGGCCCTGCCCTCTGGCCCTGGAGCTGTGGAGCCCATC	950
Qy	986	GCCAAAGAGTCTGGGATCTGGGCTTACTACGAGATCTGTACCTTCTTCAANAATGGAGCC	1045
Db	951	ACAGAGTCTTCAGGCTTCTCGGCCATATTATGAGATCTGCGAGTTCTCTGAAA - --GGAGCC	1007
Qy	1046	ACTCAGGATGGGATGCCCTCAGGAAGTGCCTTATGCCATCAGGGCAATGTGTGGGTT	1105
Db	1008	AAGATCACGGGCTCCAGGATCAGCAGGTTCCCTACCGAGTCAAGGGGAACCACTGGGTG	1067
Qy	1106	GGCTATGACAACATCAAGAGCTTCGATATTAAAGGCTCAATGGCTTAAGCAGACAACAATTT	1165
Db	1068	GGCTATGATGATGTGAAGAGTATGGAGACCAAGGTTTCAGTCTTTAAAGAAATTTAAACCTG	1127
Qy	1166	GGAGGGCCATGGTCTGGGCCATTGATCTGGATGACTTCACTGGCACTTCTTGCAACCAG	1225
Db	1128	GGAGGACCATGATCTGGTCTATTGACATGGATGACTTCACTGGCAATCTCTGCAACCAG	1187
Qy	1226	GGCAAGTTTCCCCTAATCTCCACCCCTGAAGAGGCCCTCGGC	1267
Db	1188	GGCCCTTACCTCTGTGCCAAGCAGTCAAGAGAGCCCTGGC	1229

RESULT 10

US-09-960-352-2589

; Sequence 2589, Application US/09960352

; Patent No. US20020137139A1

; GENERAL INFORMATION:

; APPLICANT: Warren, Wesley C.

; APPLICANT: Tao, Nengbing

; APPLICANT: Byatt, John C.

; APPLICANT: Mathialagan, Nagappan

; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND FAT DEPOSITION

; FILE REFERENCE: 16511.006/37-21(10298)C

; CURRENT FILING DATE: 2001-09-24

; NUMBER OF SEQ ID NOS: 15112

; SEQ ID NO 2589

; LENGTH: 415

; TYPE: DNA

; ORGANISM: Bos taurus

; OTHER INFORMATION: Clone ID: 12-LIB188-006-Q1-E1-C7

US-09-960-352-2589

Query Match 8.6%; Score 139.4; DB 10; Length 415;

Best Local Similarity 61.4%; Pred. No. 1.2e-35;

Matches 224; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

Qy 187 CACCAACTGGGCCAGTACCAGCGCCAGCGCTTGGGGCGCTTCATGCCCTGACAACATCGACCC 246

Db 12 CGCATGGGCCGGGATACCGGAGGTGATGGAGCTGCTTCCAGAGCGCCATCGACCC 71

Qy 247 CTGCGCTCTGATACCCACCTGATCTACGCCCTTTGCTGGGAGGCGAGAACACGAGATCACCAC 306

Db 72 CTTCCTGTGCACCCATGTCATCTACAGCTTTGCCAACATAAGCAACAATGAGATCGACAC 131

Qy 307 CATCGAATGGAAGATGTGACTCTCTACCAAGCTTTCAATGGCGCTGAAAAATTAAGAACAG 366

Db 132 CTGGGAGTGGAAATGACGTGACGTCTATGACACACTGAACACACTCAAGAACAGGACCC 191

Qy 367 CCAGCTCAAAACTCTCTCGCCATTGGAGCTGGAACTTCGGGACTGCCCTTTCACTGC 426

Db 192 CAACCTGAAGACCTCTCTATCTGTGGAGATGGAATTCGGTCTCAAGATTTTCCAA 251

Qy 427 CATGGTTTACTCTCTGAGAACGCCAGACTTTTCATCACCTCAGTCATCAAAATTCCTGG 486

Db 252 GATAGCTTCCAAGACCCAGAGTCGCAGGACTTTTCATCAAGTCGGTCCACCAATTTCTGG 311

Qy 487 CCAGTATGAGTTTGAAGGGCTGGACTTTGACTGGGAGTACCTCGTGGGAGCC 546

Db 312 GACCATGGCTTTGATGGACTGGACCTAGCATGGCTCTACCCCGGTGGAGACAAAGCG 371

Qy 547 TCCTC 551

Db 372 GCATC 376

RESULT 11

US-09-960-352-13050

; Sequence 13050, Application US/09960352

; Patent No. US20020137139A1

; GENERAL INFORMATION:

; APPLICANT: Warren, Wesley C.

; APPLICANT: Tao, Nengbing

; APPLICANT: Byatt, John C.

; APPLICANT: Mathialagan, Nagappan

; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND FAT DEPOSITION

; FILE REFERENCE: 16511.006/37-21(10298)C

; CURRENT FILING DATE: 2001-09-24

; NUMBER OF SEQ ID NOS: 15112

; SEQ ID NO 13050

; LENGTH: 414

; TYPE: DNA

; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 56-LIB188-019-Q1-E1-F8
US-09-960-352-13050

Query Match 8.4%; Score 136.8; DB 10; Length 414;
Best Local Similarity 62.1%; Pred. No. 8.6e-35;
Matches 216; Conservative 0; Mismatches 132; Indels 0; Gaps 0;
QY 204 ACCGGCAGCGCTGGGGCTTCATCGCTGACAAACATCGACCCCTGCTGTGACCCACC 263
DB 1 ACCGGAGGGTGTGGAGCTGCTCCAGAGCCCATGACCCCTCTGTGACCCCATG 60
QY 264 TGATCTACGCCCTTGTCTGGAGGAGCAGAACAGAGATCACACCATCGAATGGAACGATG 323
DB 61 TCATCTACAGCTTTGCAACATTAAGCAACATGAGATCGACACCTGGGAGTGAATGACG 120
QY 324 TGACTCTACCAAGCTTTCAATGGCTGAAATTAAGACAGCCAGCTGAAACCTCTCC 383
DB 121 TGACGCTCTATGACACATTTGAACACACTCAAGAACAGGAAACCCCAACCTGAAGACCCCTCC 180
QY 384 TGGCCATTGGAGCTGGAACCTTGGGAGCTGCCCTTTCACTGCGATGCTTCTACTCTCTG 443
DB 181 TATCTGTGGAGATGGAACCTTGGTCTCAAGATTTTCCAGATAGCTTCCAAACACC 240
QY 444 AGAAGCCGAGACTTTATCACTACCTAGTCATCAATTCCTGCGCCCATGATGATTTGACG 503
DB 241 GGAGTCGACGAGCTTTTCATCAAGTCGGTGCACCACTTTCTGCGGACCCATGGCTTTGATG 300
QY 504 GCGTGCATTTGACTGGGAGTACCTGGCTCTCGTGGGAGCCCTCTCTC 551
DB 301 GACTGGACCTAGCAGCTCTACCCCGGGTGGAGAGACAAAGCGGCATC 348

RESULT 12

US-09-960-352-3072
; Sequence 3072, Application US/09960352
; Patent No. US20020137139A1

; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathalagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 3072
; LENGTH: 418
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 14-LIB188-015-Q1-E1-D5
US-09-960-352-3072

Query Match 8.4%; Score 136; DB 10; Length 418;
Best Local Similarity 60.3%; Pred. No. 1.6e-34;
Matches 251; Conservative 0; Mismatches 150; Indels 15; Gaps 1;
QY 203 TACCGCCAGCGCTGGGGCTTCATGCTGTGACAAATCGACCCCTGCTGTGACCCAC 262
DB 11 TACCGGGAGGATGATGGAGCTGCGTCCAGAGCCCATGAGCCCTTCTGTGACCCCAT 70
QY 263 CTGATCTAGCCCTTCTGCGGAGCAGAACAGATCACACCATCGAATGGAACGAT 322
DB 71 GTCATCGACAGCTTTGCCAACATAAGCAACATGATGATGACACCTGGGAGTGAATGAC 130
QY 323 GTGACTCTTACCAAGCTTTCAATGGCTGAAATTAAGAACAGCCAGCTGAAACCTCTC 382
DB 131 GTGACGCTCTATGACACACTGACACACTCAAGAACAGGAAACCCCAACCTGAGACCCCTC 190
QY 383 CTGGCCATTTGGAGCTGGAGCTTGGGAGCTGCCCCCTTCTCACTGCCATGTTTCTACTCCT 442

DB 191 CTATCTGTGGAGATGGAACTTCGGTCTCTCAAGATTTTCCAAAGATAGCTTCCAAAGACT 250
QY 443 GAGAACCGGCAGACTTTTCATCCTCAGTCATCAAAATTCCTGGCCAGTATGATTTGAC 502
DB 251 CAGAGTCGCGAGGACTTTTCATCAAGTCGGTGCCACCATTTCTGCGGACCCATGCTTTGAT 310
QY 503 GGGCTGGACTTTTACTGGGAGTACCTGCTCTGCTGGGAGCCCTCTCTCAGGACAAAGCAT 562
DB 311 GGACTGGACTAGCATGCTCTACCCCGGGTGGAG-----AGACAAGCGG 355
QY 563 CTCTTCACCTGCTCTGTCGAGGAAATGCGTGAAGCTTTTCAGCAGGAGGCCAAGCA 618
DB 356 CATCTCACCACCTCTGCTCAAGGAAATGAAGGCTGAGCTTTGTAAACGGAAGCCAAAGCA 411

RESULT 13

US-09-960-352-7364
; Sequence 7364, Application US/09960352
; Patent No. US20020137139A1

; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathalagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AN
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 7364
; LENGTH: 410
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 32-LIB188-023-Q1-E1-H11
US-09-960-352-7364

Query Match 8.3%; Score 135.2; DB 10; Length 410;
Best Local Similarity 61.8%; Pred. No. 2.9e-34;
Matches 215; Conservative 0; Mismatches 133; Indels 0; Gaps 0;
QY 204 ACCGGCAGCGCTGGGGCTTCATGCTGTGACAAACATCGACCCCTGCTGTGACCCACC 263
DB 1 ACCGGAGGGTGTGGAGCTGCTCCAGAGCCCATGACCCCTTCTGTGACCCCATG 60
QY 264 TGATCTAGCGCTTTGCTGGGAGCAGAACACAGATCACACCATCGAATGGAACGATG 323
DB 61 TCATCTACAGCTTTGCCAACATTAAGCAACATGAGATCGACACCTGGGAGTGAATGACG 120
QY 324 TGACTCTCTACCAAGCTTTCAATGGCTGAAATTAAGAACAGCCAGCTGAAACCTCTCC 383
DB 121 TGACGCTCTATGACACACTGACACACTCAAGAACAGGAAACCCCAACCTGAAGACCCCTCC 180
QY 384 TGGCCATTGGAGCTGGAACCTTGGGAGCTGCCCTTCTCACTGCCATGTTTCTACTCTCTG 443
DB 181 TATCTGTGGAGATGGAACCTTGGTCTCAAGATTTTCCAAAGATAGCTTCCAAAGACC 240
QY 444 AGAAGCCGAGACTTTTCATCACCCTAGTCATCAATTCCTGCGCCAGTATGATTTGACG 503
DB 241 GGAGTCGACGAGCTTTTCATCAAGTCGGTGCACCACTTTCTGCGGACCCATGCTTTGATG 300
QY 504 GCGTGCATTTGACTGGGAGTACCTGCTGCTGGGAGCCCTCTCTC 551
DB 301 GACTGGACCTAGCATGCTCTACCCCGGGTGGAGAGACAAAGCGGCATC 348

RESULT 14

US-09-960-352-3057
; Sequence 3057, Application US/09960352
; Patent No. US20020137139A1

; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing

APPLICANT: Byatt, John C.
APPLICANT: Mathialagan, Nagappan
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
FILE OF INVENTION: MUSCLE AND FAT DEPOSITION
FILE REFERENCE: 16511.006/37-21(10298)C
CURRENT APPLICATION NUMBER: US/09/960.352
CURRENT FILING DATE: 2001-09-24
NUMBER OF SEQ ID NOS: 15112
SEQ ID NO 3057
LENGTH: 424
TYPE: DNA
ORGANISM: Bos taurus
OTHER INFORMATION: Clone ID: 14-LIB188-002-Q1-E1-D5
US-09-960-352-3057

Query Match 8.3%; Score 135.2; DB 10; Length 424;
Best Local Similarity 61.8%; Pred. No. 3e-34;
Matches 215; Conservative 0; Mismatches 133; Indels 0; Gaps 0;
Qy 204 ACCGGCCAGGCTCGGGCGCTTCATGCTGACAACTCGACCCCTGCTGTACCCACC 263
Db 1 ACCGGGAGGTGATGGAGCTGCTCCAGAGCCATCGACCCCTTCTGTGACCCATG 60
Qy 264 TGATCTAGCGCTTGTGGGAGGAGACACACAGAGATCACCACTCATGATGGAACGATG 323
Db 61 TCATCTACAGCTTGTGCAACATAAGCAACATGATGATGAGATCGACACCTGGGAGTGGATGACG 120
Qy 324 TGACTCTCTACCAAGCTTTCATGGCTGAAATAAGAACAGCAGCTGAAACTCTCC 383
Db 121 TGAGCTCTATGACACTGNACACTCAAGACAGAACCCCACTGAGACCCCTCC 180
Qy 384 TGGCCATGGAGGCTGGAAGCTGCGGAGTCCGCTTTCACATGCGCATGTTTCTACTCTG 443
Db 181 TATCTGTGGAGGATGAACTTCGTTCTCAAAGATTTTCCAAGATAGCTTCCAAGACCC 240
Qy 444 AGACCCGACAGCTTTCATCACTCAGTCATCAATTCCTGCGCCAGTATGAGTTTGACG 503
Db 241 GGAGTCGAGGACTTTCATCAAGTCGCTGCGCACCATTTCTGCGGAGCCCATGCTTTGATG 300
Qy 504 GGCTGGACTTGTAGCTGGAGTACCTGCTCTCGTGGAGCCCTCCTC 551
Db 301 GACTGGACCTAGCATGCTCTACCCCGGTGGAGACAGACGGGCATC 348

RESULT 15
US-09-960-352-14919
Sequence 14919, Application US/09960352
Patent No. US20020137139A1
GENERAL INFORMATION:
APPLICANT: Warren, Wesley C.
APPLICANT: Tao, Nengbing
APPLICANT: Byatt, John C.
APPLICANT: Mathialagan, Nagappan
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
FILE OF INVENTION: MUSCLE AND FAT DEPOSITION
FILE REFERENCE: 16511.006/37-21(10298)C
CURRENT APPLICATION NUMBER: US/09/960.352
CURRENT FILING DATE: 2001-09-24
NUMBER OF SEQ ID NOS: 15112
SEQ ID NO 14919
LENGTH: 411
TYPE: DNA
ORGANISM: Bos taurus
FEATURE:
NAME/KEY: unsure
LOCATION: (4)
OTHER INFORMATION: unsure at all n locations
OTHER INFORMATION: Clone ID: 64-LIB188-023-Q1-E1-H12
US-09-960-352-14919

Query Match 7.5%; Score 122; DB 10; Length 411;
Best Local Similarity 61.8%; Pred. No. 7.2e-30;
Matches 194; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

Qy 238 CATGACCCCTGCTCTGCTACCCACCTGATCTAGCCCTTGTGGGAGGACGAGAACAGA 297
Db 36 CATGACCCCTGCTCTGCTACCCACCTGATCTAGCCCTTGTGGGAGGACGAGAACAGA 95
Qy 298 GATCACCACCATCGAATGGAGATGATCTCTTACCAAGCTTTCAATGGCCCTGAAAAA 357
Db 96 GATCGACACCTGGGAGTGGATGACGTGACGCTCTATGACACACTGAACACACTCAAGAA 155
Qy 358 TAAGAACAGCCAGCTGAAACTCTCTGGCCATTTGGAGGCTGGAACTTCGGGACTGCCCC 417
Db 156 CAGGAACCCCAACCTGAAGACCCCTCTATCTGTTGGAGGATGGAACCTTCGGTTCTCAAAG 215
Qy 418 TTTCACTGCCATGTTTCTACTCTCTGAGAACCCGACAGCTTTTCACTCACTCAGTCATCAA 477
Db 216 ATTTTCCAAGATAGCTTCCAAGACCCGAGTCGCGAGACTTTCATCAAGTCGGTGGCCACC 275
Qy 478 ATTCCTGGCCAGTATGAGTTTGACGGGCTGGAGCTTTGACTGGGAGTACCTGCTGCTCG 537
Db 276 ATTTCTGCGGACCATGGCTTTGATGGAGTGGAGCTAGCATGGCTCTACCCCGGTGGAG 335
Qy 538 TGGGAGCCCTCCTC 551
Db 336 AGACAAGCGGCATC 349

Search completed: July 3, 2003, 10:57:23
Job time : 287.278 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2003, 04:32:16 ; Search time 2436.98 Seconds
(without alignments)
10799.277 Million cell updates/sec

Title: US-10-004-219B-2
Perfect score: 1625
Sequence: 1 gcttccagtcgtggtga.....ctgcaataaaatcagcagtc 1625

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estmu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_htc:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_htc:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: gb_gss:*
- 18: em_gss_hum:*
- 19: em_gss_inv:*
- 20: em_gss_pln:*
- 21: em_gss_vrt:*
- 22: em_gss_fun:*
- 23: em_gss_mam:*
- 24: em_gss_mus:*
- 25: em_gss_other:*
- 26: em_gss_pro:*
- 27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match Length	ID	Description
1	1047	64.4	1533 11	AK008633 Mus muscu
2	1042.2	64.1	1535 11	AK008757 Mus muscu
3	1018.4	62.7	1534 11	AK008650 Mus muscu
4	997.4	61.4	1530 11	AK007573 Mus muscu
5	772	47.5	957 13	BT760250 603045190
6	714.8	44.0	1105 14	BM923480 AGENCOURT

7	708.8	43.6	1038	13	BI511797	BI511797 603042049
8	681.8	42.0	892	14	BQ231098	BQ231098 AGENCOURT
9	653.6	40.2	872	14	BQ231064	BQ231064 AGENCOURT
10	649.8	40.0	838	12	BG867815	BG867815 602786336
11	641.2	39.5	824	13	BI518221	BI518221 603042049
12	630	38.8	855	12	BG872136	BG872136 602792974
13	629	38.7	837	12	BG868741	BG868741 602787570
14	626	38.5	847	12	BG866775	BG866775 602786605
15	623.2	38.4	809	12	BG866417	BG866417 602785352
16	622.4	38.3	836	12	BG869655	BG869655 602789716
17	620.2	38.2	1347	14	BM924663	BM924663 AGENCOURT
18	619.4	38.1	833	12	BG871337	BG871337 602790509
19	617.8	38.0	880	12	BG871468	BG871468 602790678
20	617.4	38.0	863	12	BG872850	BG872850 602793973
21	614.6	37.8	894	12	BG869958	BG869958 602789826
22	611.2	37.6	874	12	BG865756	BG865756 602784345
23	610.6	37.6	804	12	BG872805	BG872805 602793325
24	608.4	37.4	839	12	BG870376	BG870376 602791289
25	607.6	37.4	856	12	BG870147	BG870147 602793390
26	607.4	37.4	837	12	BG865578	BG865578 602783726
27	607.4	37.4	845	12	BG865848	BG865848 602788264
28	606.6	37.3	845	12	BG867145	BG867145 602786759
29	603.2	37.1	769	12	BG868463	BG868463 602785844
30	603	37.1	822	12	BG870367	BG870367 602791260
31	602.6	37.1	799	12	BG872144	BG872144 602792982
32	602	37.0	763	12	BG872647	BG872647 602793705
33	601.8	37.0	794	12	BG869492	BG869492 602789190
34	601.6	37.0	764	12	BG869304	BG869304 602788819
35	600	36.9	848	12	BG869143	BG869143 602789020
36	598.2	36.8	939	14	BQ219632	BQ219632 AGENCOURT
37	597.6	36.8	831	12	BG867708	BG867708 602786811
38	597.2	36.8	838	12	BG870392	BG870392 602791292
39	596.8	36.7	886	12	BG870840	BG870840 602791891
40	596.2	36.7	882	12	BG873488	BG873488 602791730
41	595.8	36.7	948	13	BI664452	BI664452 603290028
42	595.6	36.7	817	12	BG873398	BG873398 602794380
43	595.4	36.6	845	12	BG865076	BG865076 602784229
44	595.2	36.6	631	13	BI761276	BI761276 603044029
45	595	36.6	1155	12	BG867473	BG867473 602788579

ALIGNMENTS

RESULT 1	AK008633	1533 bp	mRNA	linear	HTC 19-JAN-2002
LOCUS	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:2200003E03:chitinase, acidic, full insert sequence.				
DEFINITION	AK008633				
ACCESSION	AK008633.1	GI:12842941			
VERSION	HTC: CAP trapper				
KEYWORDS	Mus musculus (strain:C57BL/6J) adult male stomach cDNA to mRNA, clone:lib:RIKEN full-length enriched mouse cDNA library				
SOURCE	clone:2200003E03.				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.				
AUTHORS	1				
TITLE	Carninci, P. and Hayashizaki, Y.				
JOURNAL	High-efficiency full-length cDNA cloning				
MEDLINE	Meth. Enzymol. 303, 19-44 (1999)				
PUBMED	99279253				
REFERENCE	10349636				
AUTHORS	2				
TITLE	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
JOURNAL	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
MEDLINE	Genome Res. 10 (10), 1617-1630 (2000)				
PUBMED	20499374				
REFERENCE	11042159				

[illegible]

RESULT 2				
AK008757				
LOCUS	AK008757	1535 bp	mRNA	linear
DEFINITION	Mus musculus adult male stomach cDNA, RIXEN full-length enriched library, clone:2210020N23:chitinase, acidic, full insert sequence.			

ACCESSION	AK008757
VERSION	AK008757.1
KEYWORDS	GI:12843148
SSOURCE	HTC; CAP trapper.
	Mus musculus (strain:C57BL/6J) adult male stomach cDNA to mRNA,
	clone_lib:RIKEN full-length enriched mouse cDNA library
	clone_id:2210020N23.
ORGANISM	Mus musculus
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE	
AUTHORS	Carninci, P. and Hayashizaki, Y.
TITLE	High-efficiency full-length cDNA cloning
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)
MEDLINE	99279253
PUBMED	10349636
REFERENCE	
AUTHORS	2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
	Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to
JOURNAL	prepare full-length cDNA libraries for rapid discovery of new genes
MEDLINE	Genome Res. 10 (10), 1617-1630 (2000)
PUBMED	20499374
REFERENCE	
AUTHORS	3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
	Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
TITLE	Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishide, T., Kashiwagi, K.,
JOURNAL	Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
MEDLINE	Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
PUBMED	Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
REFERENCE	Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
AUTHORS	RIKEN integrated sequence analysis (RISA) system--384-format
TITLE	sequencing pipeline with 384 multicapillary sequencer
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE	20530913
PUBMED	11076861
REFERENCE	
AUTHORS	4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
	Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
TITLE	Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
JOURNAL	Saito, T., Okazaki, Y., Gotohori, T., Bono, H., Kasukawa, T., Saito, R.,
MEDLINE	Kadota, K., Matsuda, H., Asahiner, M., Batalov, S., Casavant, T.,
PUBMED	Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
REFERENCE	Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G.,
AUTHORS	Quackenbush, J., Schirml, L. M., Stauber, F., Suzuki, R., Tomita, M.,
	Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
TITLE	Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
JOURNAL	Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C.,
MEDLINE	Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
PUBMED	Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,
REFERENCE	Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
AUTHORS	Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
	Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H.,
TITLE	Toyo-oka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L.,
JOURNAL	Yinshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S.,
MEDLINE	and Hayashizaki, Y.
PUBMED	Functional annotation of a full-length mouse cDNA collection
REFERENCE	Nature 409 (6821), 685-690 (2001)
AUTHORS	21085660
	11217851
TITLE	5 (bases 1 to 1535)
JOURNAL	Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,
MEDLINE	Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C.,
PUBMED	Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T.,
REFERENCE	Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hirooka, T., Hori, F.,
AUTHORS	Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T.,
	Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S.,
TITLE	Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K.,
JOURNAL	Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J.,
MEDLINE	Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
PUBMED	Schirml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T.,
REFERENCE	Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,
AUTHORS	Tajima, Y., Toya, T., Yamamura, T., Yamanaka, I.,

161 TCTGCTTACAGCTGACATCTACTTCCACCACTGGCCAGTACCGCCAGGCGCTGGG 220
162 TCTGCTTACAGCTGACATCTACTTCCACCACTGGCCAGTACCGCCAGGCGCTGGG 220
74 TCTGCTTACATCTGATATCTATTTCCACCACTGGCCAGTATCGCCAGGCTGGG 133
221 CGCTTACAGCTGACATCTGACCTGCTCTGTACCACTGATCTACGCTTTGCT 280
134 AGCTTCAAGCTGATGACATTAACCTGCTGTGTACTCACCTGATCTATGCTTTGCT 193
281 GGGAGCAGACACAGAGATCACCACCACTGAATGGAGCAGTGTACTCTTACCAGCT 340
194 GGGAGCAGACACAGATGATGATCACCACCACTGAATGGAGTGTACTCTTAAAGCT 253
341 TTCAATGGCTGAAAATAAGACAGCAGCTGAAAACCTCTCTGGCCATTTGGAGGTGG 400
254 TTCAATGATTTGAACACAGACAGCAAACTGAAACCTCTCTGGCAATTTGGAGGTGG 313
401 AACTTGGGAGCTGCTTCTTCACTGCGATGGTTTCTACTCTCTGAGAACCCGACAGCTTC 460
314 AACTTTGGAATGCTCTCTTCACTTACCATGCTTCTTCCACTTCTCAGAACCCGACAGCTTC 373
461 ATCACTCTCAGTATCAATCTCTGCGCCAGTATGAGTTTTCAGCGGCTGGAGCTTGCAGTGG 520
374 ATTACCTCAGTATCAATCTTCTGCTCAGTATGGTTTGTAGACTGGAGCTGGAGCTGG 433
521 GAGTACCTGCTCTCTGCTGGAGCCCTCTCTCAGGACAGCATCTTCTACTGTCTGGTG 580
434 GAATACCCAGGCTCAGCTGGAGCCCTCTCTCAGGACAGCATCTTCTACTGTCTGGTG 493
581 CAGGAATGCTGAGCTTTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 640
494 AAGGAATGCTGAGCTTTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 553
641 GTCACTGCTCAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 700
554 GTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 613
701 CTGTACAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 760
614 CTTTAAAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 673
761 GGTACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 820
674 GGTACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 732
821 TACCTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 880
733 TACCTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 792
881 CTCACTGCT 940
793 CTCACTGCT 851
941 GGAATGCT 1000
852 GGAATGCT 911
1001 ATCTGGCTTACTACGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1060
912 TTCTGGGCTTACTATGAGATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 971
1061 GCGCTCAGGAAGTCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1120
972 GCGCTCAGGAAGTCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1031
1121 AAGAGCTTCAATATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1180
1032 AAGAGCTTCAATATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1091
1181 TGGGCTTATGCT 1240
1092 TGGGCTTATGCT 1151
1241 ATCTCCACCTGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1300

1152 ACTTCTACTTTGAACAAAGCCCTTGGCATATCCACTGAAGTTGCACAGCTCTCTGACGTG 1211
1301 CCATTGAGCCATTAATCTGCTGCCAGTGGCAGCGGGAACGGGAGCGGAGTAGCAGC 1360
1212 CCTTCCGAGCCAGTGAATCTACTCTCTCC-----AGGAAGTGGAGTGGGGTGGGAGC 1262
1361 TCTGGAGGAGCTCGGAGGAGCTGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1420
1263 TCCGAGGAGCTCTGGAGCAGTGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1322
1421 GTGCAAAATAACAGAAATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1480
1323 GTGCAAGATCAGCAAAATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1382
1481 TGCAGGCGGCT 1540
1383 TGTCAAGCAGGCT 1442

RESULT 4
AK007573
LOCUS
DEFINITION
Mus musculus 10 day old male pancreas cDNA, RIKEN full-length
enriched library, clone:1810021D23:chitinase, acidic, full insert
sequence.
ACCESSION
AK007573
VERSION
AK007573.1 GI:12841201
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (strain:C57BL/6J) 10 day old male pancreas cDNA to
mRNA, clone_lib:RIKEN full-length enriched mouse cDNA library
clone:1810021D23.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
MEDLINE
10349636
2
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20493374
MEDLINE
11042159
3
Shibata,K., Itoh,M., Aizawa,K., Nagaoaka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multiplexed capillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
MEDLINE
11076861
4
Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
Arakawa,T., Harai,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,
Saito,T., Okazaki,Y., Gotohori,T., Bono,H., Kasukawa,T., Saito,R.,
Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G.,
Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,
Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Anoh,H.,
Baldarelli,R., Barsh,G., Blake,J., Botfelli,D., Bojunga,N.,


```
QY 1178 GTCTGGCCCATGATCGGATGACTTCACCTGGCAGCTTCCTGCAACAGGCGCAAGTTTCCC 1237
Db 1086 ATCTGGCCCATGACCTTGATGACTTCACCTGGCTCTTCTGTCATCAGGGAATTTCT 1145
QY 1238 CTAATCTCCACCTGAAGAGGCGCTCGGCTCGCAGAGTCAGTTGCGAGGCTCAGCT 1297
Db 1146 CTGACTTCTACTTTGAACAAGCCCTGGCATATCCACTGAAGGTTGACAGCTCCTGAC 1205
QY 1298 CAGCCCATGAGCAATTAATGCTGCTCCAGTGGCAGCGGGAACGGAGCGGAGTAGC 1357
Db 1206 GTCCCTCCGAGCGAGTACTCTCTCC-----AGGAAGTGGAGTGGGGTGA 1256
QY 1358 AGCTTGGAGGAGCTCGGAGGAGTGGATTCGTCTGTAGAGCCCAACGCCCTCTAC 1417
Db 1257 AGCTCGGAGGAGCTCTGGAGCAGTGGATTCGTGCCACAAAGCAGATGCGCTCTAC 1316
QY 1418 CCCGTGGCAATTAACAGAAATGCTTCTGGCAGTGGTGAATGGAGTACGATACCGAG 1477
Db 1317 CCGTGGCAGATACAGAAATGCTTTTGGCAGTGCATCAATGGAATCAATACCGAG 1376
QY 1478 AACTGGCAGGCGGCTTCTTCGACACAGCTGTGATGCTGCAACTGGGCGATAAAC 1537
Db 1377 CATGTGCAAGCGGCTGTTTGTATACCACTGTAAATGCTGCAACTGGCCATGACC 1436
QY 1538 TGA 1540
Db 1437 TAA.1439

RESULT 5
BI760250
LOCUS 603045190F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5185486 5',
DEFINITION mRNA sequence.
ACCESSION BI760250
VERSION BI760250.1 GI:15751828
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 957)
NIH-MGC http://mgc.ncl.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs@email.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Cloned and Sequenced by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11463 Row: c Column: 23
High quality sequence stop: 780.
Location/Qualifiers
1..957
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5185486"
/lab_host="NIH_MGC_116"
/notes="Organ: pooled colon, kidney, stomach; Vector:
pcMV-SPT6; Site:1: NotI; Site:2: EcoRV (destroyed); RNA
source anonymous pool of 3 colons, age 26 yo male, 49 yo
female, 71 yo male colon; 46 yo male kidney, and pool of 2
females, 62 yo male and 70 yo female. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.4 kb.
Insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
```

```
BASE COUNT 222 a 277 c 227 g 231 t
ORIGIN
Query Match 47.5%; Score 772; DB 13; Length 957;
Best Local Similarity 94.3%; Pred. No. 5.3e-204;
Matches 877; Conservative 0; Mismatches 45; Indels 8; Gaps 7;
339 CTTTCAATGCGCTGAAATAAAGACAGCAGCTGAAATCTCTCTGGCCATTGGAGGCT 398
Db 25 CTTTACTGAAATTTGACGCTCGGAACAGCAGCTGAAATCTCTCTGGCCATTGGAGGCT 84
QY 399 GGAATTCGGAGTGCCTTTTCACTGCCATGCTTTTCTACTCTGAGAACCCGAGACTT 458
Db 85 GGAATTCGGAGTGCCTTTTCACTGCCATGCTTTTCTACTCTGAGAACCCGAGACTT 144
QY 459 TCATCACCCTCAGTCATCAATTTCTTGGCCAGTATGATTTGACGGCTGGACTTTGACT 518
Db 145 TCATCACCCTCAGTCATCAATTTCTTGGCCAGTATGATTTGACGGCTGGACTTTGACT 204
QY 519 GGGAGTACCTTGGCTCTCGTGGAGCGCTCTCTCAGGACAAGCATCTCTTCACTGTCTGG 578
Db 205 GGGAGTACCTTGGCTCTCGTGGAGCGCTCTCTCAGGACAAGCATCTCTTCACTGTCTGG 264
QY 579 TGCAGGAAATGCGTGAAGCTTTTGGAGCAGAGGCCAAGCAGATCAACAAGCCAGGCTGA 638
Db 265 TGCAGGAAATGCGTGAAGCTTTTGGAGCAGAGGCCAAGCAGATCAACAAGCCAGGCTGA 324
QY 639 TGGTCACTGCTGAGTACCTTGGCATCTCCAAATATCCAGTCTGCTATGAGATCCCC 698
Db 325 TGGTCACTGCTGAGTACCTTGGCATCTCCAAATATCCAGTCTGCTATGAGATCCCC 384
QY 699 AACTGTACAGTACCTTGGACTACATCCATGTCATGACCTACGACCTTCCATGGCTCTCTGG 758
Db 385 AACTGTACAGTACCTTGGACTACATCCATGTCATGACCTACGACCTTCCATGGCTCTCTGG 444
QY 759 AGGGCTACACTGGAGAGAACAGCCCTCTTACAATAACCCGACTGACACCGCAGCAACG 818
Db 445 AGGGCTACACTGGAGAGAACAGCCCTCTTACAATAACCCGACTGACACCGCAGCAACG 504
QY 819 CCTACCTCAATGTGGATTATGTCATGAATCTGGAAGGACATGAGCAGCAGCTGAGA 878
Db 505 CCTACCTCAATGTGGATTATGTCATGAATCTGGAAGGACATGAGCAGCAGCTGAGA 564
QY 879 AGCTCATCTGTGGATTCCCTTACCTATGGACACAATTTCACTGAGCAACCCCTCCAACA 938
Db 565 AGCTCATCTGTGGATTCCCTTACCTATGGACACAATTTCACTGAGCAACCCCTCCAACA 624
QY 939 CTGGAATTTGGTGGCCCGCCACTCTGTGTGCTGGTCTCTGCGGCCCTATGCCAAGAGTCTG 998
Db 625 CTGGAATTTGGTGGCCCGCCACTCTGTGTGCTGGTCTCTGCGGCCCTATGCCAAGAGTCTG 684
QY 999 GGATCTGGGCTTACTACGAGATCTGTACCTTCTCTGAA-AAATGAGCCACTCAGGATGG 1057
Db 685 GGATCTGGGCTTACTACGAGATCTGTACCTTCTCTGAA-AAATGAGCCACTCAGGATGG 744
QY 1058 GATGCCCCCTC-AGGAAGTGCCTTATGCCTA-TCAGGGCAATGTGGGTGGCTAT--GA 1113
Db 745 GATGCCCCCTCAGGAAGTGCCTTATGCCTATTCAGGGCAGACAGTGTGGTGGCTATGAC 804
QY 1114 CAACATCAAGAGC-TTCGATATTAAAGGTCATATGCG-TTAAGCACAAATAATTGGAGGC 1171
Db 805 CAACATCAAGAGCTTTTCGATATTAAAGGTCATATGCGTTTAAAGCACAACCAATTTGAGGC 864
QY 1172 GCATGGTCTGGCCCAATTCATCT-GGATGACTTCACTGGCAGCTTTCTGCAACCCAGGCGAA 1230
Db 865 GCATGGTCTGGCCCAATTCATCTGCGGATTAACCTTCCACTGGGCACTTCTTCAACCCAGGCGAA 924
QY 1231 GTTTCCCTTAATCTCCACCCTCGAAGAAGCG 1260
Db 925 GTTTCCCTTAATTTCCACCCTGTAACAGCG 954
```



```

RESULT 6
BM923480      1105 bp      mRNA      linear      EST 12-MAR-2002
LOCUS      AGENCOURT_6623879 NTH_MGC_116 Homo sapiens cDNA clone IMAGE:5759117
DEFINITION      5', mRNA sequence.
ACCESSION      BM923480
VERSION      BM923480.1 GI:19373859
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE      NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS      National Institutes of Health, Mammalian Gene Collection (MGC).
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-remail.nih.gov
              Tissue Procurement: Life Technologies, Inc.
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Agencourt Bioscience Corporation
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LLAM12804 row: a column: 06
              High quality sequence stop: 668.
              Location/Qualifiers
                1..1105
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:5759117"
                  /clone_lib="NIH_MGC_116"
                  /lab_host="DH10B"
                  /note="Organ: pooled colon, kidney, stomach; Vector:
                  pCMV-SPORT6; Site.1: NotI; Site.2: EcoRV (destroyed); RNA
                  source anonymous pool of 3 colons, age 26 yo male, 49 yo
                  female, 71 yo male colon; 46 yo male kidney, and pool of 2
                  stomachs, 62 yo male and 70 yo female. Library is
                  oligo-dT primed and directionally cloned (EcoRV site is
                  destroyed upon cloning). Average insert size 1.4 kb,
                  insert size range 1-3 kb. Library is normalized and
                  enriched for full-length clones and was constructed by C.
                  Gruber (Invitrogen). Research Genetics tracking code
                  023. Note: this is a NIH_MGC Library."
BASE COUNT      270 a 318 c 285 g 231 t 1 others
ORIGIN
Query Match      44.0%; Score 714.8; DB 14; Length 1105;
Best Local Similarity 97.3%; Pred. No. 5.5e-188;
Matches 759; Conservative 0; Mismatches 17; Indels 4; Gaps 3;

QY 584 GAATGCGTGAAGCTTTTGAGCAGAGGCCAGCAGATCAACAGCCAGGCTGATGGTC 643
      |||||||
Db 121 GAATGCGTGAAGCTTTTGAGCAGAGGCCAGCAGATCAACAGCCAGGCTGATGGTC 180
      |||||||
QY 644 ACTGCTGCAGTAGCTGCTGGCATCTCCAAATATCCAGTCTGGCTATGAGATCCCAACTG 703
      |||||||
Db 181 ACTGCTGCAGTAGCTGCTGGCATCTCCAAATATCCAGTCTGGCTATGAGATCCCAACTG 240
      |||||||
QY 704 TCACAGTACCTGGACTACATCATGTCATGACCTACGACCTCCATGGCTCTCTGGGAGGCG 763
      |||||||
Db 241 TCACAGTACCTGGACTACATCATGTCATGACCTACGACCTCCATGGCTCTCTGGGAGGCG 300
      |||||||
QY 764 TACACTGGAGAACAGCCCCCTCTACAAATACCGACTGACACCGGACGACCGCTTAC 823
      |||||||
Db 301 TACACTGGAGAACAGCCCCCTCTACAAATACCGACTGACACCGGACGACCGCTTAC 360
      |||||||
QY 824 CTCATATGCGATTATGTCATGAATGAGGAGCAATGGAGCACCAGCTGAGAGCTC 883
      |||||||
Db 361 CTCATATGCGATTATGTCATGAATGAGGAGCAATGGAGCACCAGCTGAGAGCTC 420
      |||||||
QY 884 ATCGTTGGATTCCCTACCTATATGGACACAACTTTCATCCTGAGCAACCCCTCCAACACTGGA 943
      |||||||
Db 421 ATCGTTGGATTCCCTACCTATATGGACACAACTTTCATCCTGAGCAACCCCTCCAACACTGGA 480
      |||||||
QY 944 ATTGGTCCCCCACCCTCTGGTCTGCTGCTGGCCCTATGCGCAAGAGTCTGGGATC 1003
      |||||||
Db 481 ATTGGTCCCCCACCCTCTGGTCTGCTGCTGGCCCTATGCGCAAGAGTCTGGGATC 540
      |||||||
QY 1004 TGGGCTTACTACGAGATCTGTACCTTCTCTGAAAATGAGGCACATCAGGATGGATGCC 1063
      |||||||
Db 541 TGGGCTTACTACGAGATCTGTACCTTCTCTGAAAATGAGGCACATCAGGATGGATGCC 600
      |||||||
QY 1064 COTCAGGAAGTGCCTTATGCTCCTATCAGGCAATGTGTGGTGGCTATGACAACTCAAG 1123
      |||||||
Db 601 COTCAGGAAGTGCCTTATGCTCCTATCAGGCAATGTGTGGTGGCTATGACAACTCAAG 660
      |||||||
QY 1124 AGCTTCGATATTAAGGCTCAATGGCTTAAGCACACAAATTTGGAGGCGCCATGCTGG 1183
      |||||||
Db 661 AGCTTCGATATTAAGGCTCAATGGCTTAAGCACACAAATTTGGAGGCGCCATGCTGG 720
      |||||||
QY 1184 GCATTGATCTGGATGACTTCACTGGCACTTTCTCAACCCAGGCAAGTTTCCCTAATC 1243
      |||||||
Db 721 GCATTGATCTGGATGACTTCACTGGCACTTTCTCG-ACCAGGCAAGTTTCCCTAATC 779
      |||||||
QY 1244 TCCACCC--TGAAGAAGGCCCTCGGCCCTGCAGAGTGCAAGTTGCA-CGGCTCCAGCTCAG 1300
      |||||||
Db 780 TCCACCCCTGAAAGAAGGCCCTCGGCCCTGCAGAGTGCAAGTTGCAAGCTCCAGCTCAG 839
      |||||||
QY 1301 CCCATTGAGCCCAATAACTGCTGCCAGTGCAGCGGACGGGACGGGAGTACGAGC 1360
      |||||||
Db 840 CCCATTGAGCCCAATAACTGCTGCCAGTGCAGCGGACGGGACGGGAGTACGAGTAC 899
      |||||||

RESULT 7
BI517797      1038 bp      mRNA      linear      EST 29-AUG-2001
LOCUS      603042049F1 NTH_MGC_116 Homo sapiens cDNA clone IMAGE:5182357 5',
DEFINITION      mRNA sequence.
ACCESSION      BI517797
VERSION      BI517797.1 GI:15342589
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE      NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-remail.nih.gov
              Tissue Procurement: Life Technologies, Inc.
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LLAM11455 row: a column: 14
              High quality sequence start: 3
              High quality sequence stop: 788.
              Location/Qualifiers
                1..1038
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:5182357"
                  /clone_lib="NIH_MGC_116"
                  /lab_host="DH10B"
                  /note="Organ: pooled colon, kidney, stomach; Vector:
                  pCMV-SPORT6; Site.1: NotI; Site.2: EcoRV (destroyed); RNA
                  source anonymous pool of 3 colons, age 26 yo male, 49 yo
                  female, 71 yo male colon; 46 yo male kidney, and pool of 2
                  stomachs, 62 yo male and 70 yo female. Library is
                  oligo-dT primed and directionally cloned (EcoRV site is

```

destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 023. Note: this is a NIH_MGC Library."

BASE COUNT	273 a	275 c	293 g	197 t	
ORIGIN					
Query Match	43.6%	Score	708.8	DB 13	Length 1038;
Best Local Similarity	93.8%	Pred. No.	2.5e-186;		
Matches 784; Conservative	0;	Mismatches	42;	Indels	10; Gaps 4;
QY	769	TGGAGAGACAGCCCCCTCTACAAATACCCGACTGACACCGCAGCAACGCCCTACCTCAA	828		
DB	1	TGGAGAGACAGCCCCCTCTACAAATACCCGACTGACACCGCAGCAACGCCCTACCTCAA	60		
QY	829	TGTGGATTATGTCTATGAACCTACTTGGAGGACAAATGGAGCACCAGCTCAGAAAGCTCATCGT	888		
DB	61	TGTGGATTATGTCTATGAACCTACTTGGAGGACAAATGGAGCACCAGCTCAGAAAGCTCATCGT	120		
QY	889	TGGATTCCTTACCTATGAGACACAACCTTCATCTTGAGCAACCCCTCCAACTCTGGAATTTGG	948		
DB	121	TGGATTCCTTACCTATGAGACACAACCTTCATCTTGAGCAACCCCTCCAACTCTGGAATTTGG	180		
QY	949	TGCCCCACCTCTGGTGCTGGTCTGGTCTGGGCCCTATGCCAAGAGTCTGGGATCTGGGC	1008		
DB	181	TGCCCCACCTCTGGTGCTGGTCTGGTCTGGGCCCTATGCCAAGAGTCTGGGATCTGGGC	240		
QY	1009	TTACTAGGAGATCTGTACCTTCTCTGAAAAATGGAGCCACCTCAGGGATGGGATGCCCTCA	1068		
DB	241	TTACTAGGAGATCTGTACCTTCTCTGAAAAATGGAGCCACCTCAGGGATGGGATGCCCTCA	300		
QY	1069	GGAAGTGCCCTTATGCGCTATCAGGCGCAATGTGTGGTGTGGCTATGACAACTCATCAGAGCTT	1128		
DB	301	GGAAGTGCCCTTATGCGCTATCAGGCGCAATGTGTGGTGTGGCTATGACAACTCATCAGAGCTT	360		
QY	1129	CGATATTAAAGGCTCAATGGCTTTAAGCAACAACAAATTTGGAGGCCCATGTCTGGGGCAT	1188		
DB	361	CGATATTAAAGGCTCAATGGCTTTAAGCAACAACAAATTTGGAGGCCCATGTCTGGGGCAT	420		
QY	1189	TGATCTGGATGACTTCACTGGCACTTTCTTGCAACACGAGGCAAGTTTCCCTTAATCTCCAC	1248		
DB	421	TGATCTGGATGACTTCACTGGCACTTTCTTGCAACACGAGGCAAGTTTCCCTTAATCTCCAC	480		
QY	1249	CCTTGAGAGAGGCCCTTGGCCCTGCAGAGTGCAGAGTTGCACGGCTCCAGCTCAGGCCATTTGA	1308		
DB	481	CCTTGAGAGAGGCCCTTGGCCCTGCAGAGTGCAGAGTTGCACGGCTCCAGCTCAGGCCATTTGA	540		
QY	1309	GCCAAATACTGCTGCTCCCAAGTGGCAGCGGGAACGGGAGCGGAGTAGCAGCTCTGGAGG	1368		
DB	541	GCCAAATACTGCTGCTCCCAAGTGGCAGCGGGAACGGGAGCGGAGTAGCAGCTCTGGAGG	600		
QY	1369	CAGCTCGGGAGGCAAGTGGATTTCT - GTGCTGTGCAGAGCCCAACGCCCTCTACCCCTGGCA	1426		
DB	601	CAGCTCGGGAGGCAAGTGGATTTCT - GTGCTGTGCAGAGCCCAACGCCCTCTACCCCTGGCA	660		
QY	1427	AATA - - - - ACAGAAATGCCTTCTGGCACTGCGTGAATGAGTCAAGTACAG - - - - CAGAA	1479		
DB	661	CAAAATTACCAAGATGCGCTTCTGGGCACCTGCGTGAATGAGTCAAGTACAGTACCGAGCAAC	720		
QY	1480	CTGCCAGGCGGGCTTGT - CTTCCAGACACAGCTGTGATTCCTGCAACTGGGCAATAAACCT	1538		
DB	721	TTGCCAGGCGGGCTTGT - CTTCCAGACACAGCTGTGATTCCTGCAACTGGGCAATAAACCT	780		
QY	1539	GACCTGGCTATATTCCTTAGAGTTCAGTCTCTTTTGGCTTAGGACATGTTGCC	1594		
DB	781	GACCTGGCTATATTCCTTAGAGTTCAGTCTCTTTTGGCTTAGGACATGTTGCC	836		

IMAGE:6051580 5', mRNA sequence.

BQ231098

BQ231098.1 GI:20412498

EST.

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 892)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: L1AM13305 row: k column: 05

High quality sequence stop: 772.

Location/Qualifiers

1..892

/organism="Mus musculus"

/db_xref="taxon:10090"

/clone="IMAGE:6051580"

/clone_lib="NCI_CGAP_St1"

/lab_host="DH10B (TI-resistant)"

/note="Organ: stomach; Vector: pCMV-SPOPT6; Site_1: Sali; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.77 kb. Library constructed by Life Technologies. Note: this is a NCI_CGAP Library."

219 a 226 c 216 g 230 t 1 others

BASE COUNT

ORIGIN

Query Match	42.0%	Score 681.8;	DB 14;	Length 892;
Best Local Similarity	85.6%;	Pred. No. 7.7e-179;		
Matches 758;	Conservative 0;	Mismatches 128;	Indels 0;	Gaps 0;
QY	132	TTGTCCTTATACATGAATTTTCAGCTGGGCTCTGCCTACCAGCTGCACATGCTACTTCACCA	191	
Db	1	TTGGCTCTTCGTGTAATGCTCAGCTGGGCTCTGCCTACATCTGATATGCTATTTACCA	60	
QY	192	ACTGGGCCCAAGTACGGGCCAGGCCCTTTCATGCTGTGACACATCGACCCCTGCC	251	
Db	61	ACTGGGCCCAAGTATCGGCCAGGCTGCGGGAGCTTCAAGCCTGATGACATTTAACCCCTGCC	120	
QY	252	TCGTGTACCCACCTGATCTACGCCCTTTGCTGGGAGGAGAACACGAGATCACCACCATCG	311	
Db	121	TGTGTACTACCTGATCTATGCCCTTTGCTGGGATGCAGACATGAGATCACCACCATAG	180	
QY	312	AATGGAACGATGTGACTCTCTACCAAGCTTTTCATATGGCCTTGAAAAATAAGACAGCCAGC	371	
Db	181	AATGGGAATGATGTTACTCTCTATAAAGCTTTTCATGACTTTGAAAAACAGGAACAGCAAAAC	240	
QY	372	TGAAAACCTCCCGGCATTTGGAGGCTTGGAACTTCGGGACTGCCCTTTTCACATGCCATGG	431	
Db	241	TGAAAACCTCCCTGGCAATTTGGAGGCTTGGAACTTGGAACTGCTCTCTTCACATACCATGG	300	
QY	432	TTTCTTACTCCTGAGAACCCGCGAGACTTTTCATCACCTCAGTCATCAAAATTCCTCGGCCAGT	491	
Db	301	TTTCCACTCTCTCAGAACCGCGACACCTTTCATTAACCTCAGTCATCAAAATTTCTCGCTCAGT	360	
QY	492	ATCAGTTTCACGGGCTGGACTTTTACTTGGGAGTACCTTGCTCTCTCGTGGGAGCCCTCCCTC	551	
Db	361	ATGGGTTTGATGGACTTGGACCTGACTTGGGAATACCCAGGCTCACGTGGGAGCCCTCCCTC	420	
QY	552	AGGACAAGCATCTCTTCACTGTCTCGTGCAGGAAATCGGTGAAGCTTTTGGACGAGGAGG	611	
Db	421	AGGACAAGCATCTCTTCACTGTCTCGTGCAGGAAATCGGTGAAGCTTTTGGACGAGGAGG	480	

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (Bases 1 to 838)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: rcspbs-re@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LHAM0816 row: g column: 09
High quality sequence stop: 824.
Location/Qualifiers
1. 838
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone_image="IMAGE:491254"
/clone_lib="NCI_CGAP_SG2"
/lab_host="DH10B (T1 phage-resistant)"
Note="Organ: salivary gland; Vector: pCMV-SPORT6; Site:1:
Not; Site:2: SalI; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 209 a 211 c 198 g 220 t
ORIGIN
Query Match 40.0%; Score 649.8; DB 12; Length 838;
Best Local Similarity 87.0%; Pred. No. 6.3e-170;
Matches 714; Conservative 0; Mismatches 107; Indels 0; Gaps 0;
Qy 137 CTTATACGTAATTTGCAGCTCGGCTCGCTACAGCTGACATGCTACTTCCACCACTGG 196
Db 4 CTTCTGCTGAATGCTCAGCTCGGGCTCGCTACAACTCTGATATGCTATTTCCACCACTGG 63
Qy 197 GCCAGTACCGGCCAGCGCTGGGGCGCTTCATGCTGACACATCGACCCCTGCCTCTGT 256
Db 64 GCCAGTATCGGCCAGCTCTGGGAGCTTCNAGCCTGATGACATTACCCTGCTGTGT 123
Qy 257 ACCCACCTGATCTACGCTTTTGTGGGAGGAGAACACGAGATCACACCATCGAATGG 316
Db 124 ACTCACCTGATCTATGCTTGTGTGGATGACAGAACAAATGAGATCACCCACCATAGATGG 183
Qy 317 AACGATGTACTCTCTACCAAGCTTTCAATGGCTGAAAAATAGACAGCCAGCTGAAA 376
Db 184 AATGATGTTACTCTCTAAGCTTTCAATGACTTGAACAGAGAACAGCAAACTGAAA 243
Qy 377 ACTCTCTGCCCATTTGGAGCTGGAACCTCGGGACTCGCCCTTTCACTGCCATGTTTCT 436
Db 244 ACCCTCTGCAATTTGGAGGCTGGAACCTTTGGAACTGCTCCTTTCACTACCATGTTTCC 303
Qy 437 ACTCTGAGAACGCCAGACATTTTCATCACCTCAGTCATCAAAATTCCTTGGCGGCAATGAG 496
Db 304 ACTTCTCAGAACGCCAGACCTTCATTTACCTCAGTCATCAAAATTTCTCGCTCAGTATGG 363
Qy 497 TTTGACGGCTGGACTTTGACTGGGAGTACCCCTGGCTCTGTTGGAGCCCTCCCTCAGAC 556
Db 364 TTTGATGGACTGGACTGGACTGGGATACCCAGGCTCACGTGGGAGCCCTCTCTCAGAC 423
Qy 557 AAGCATCTTTCACTGTCTCTGGTGCAGGAAATGCGTGAAGCTTTTGTAGCAGGAGCCCAAG 616
Db 424 AAGCATCTTTCACTGTCTCTGGTGAAGGAAATGCGTGAAGCTTTTGTAGCAGGAGCTATT 483
Qy 617 CAGATCAACAGCCAGGCTGATGGTCACTGCTCAGTACGTGCTGGCATCTCCAATATC 676
Db 484 GAGAGCAGCAGCCAGACTGATGGTTACTTGTGCTAGCTGGTGGGATTTCCAAATC 543
Qy 677 CAGTCTGGCTATGAGATCCCCCAACTCTCACAGTACCTGGACTACATCCATGTGATGACC 736

Db 544 CAGGCTGGCTATGAGATCCCTCGAACTTTCCTTAAGTACTCCTGGATTTCATCCATGCATGACA 603

QY 737 TACGACCTCCATGTGCTCCTGGAGGGGTACACCTGGAGAACACGCCCTCTTACAATAAC 796
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 604 TATGACCTCCATGTGCTCCTGGAGGGGTACACCTGGGAGAATAGTCTCTTTACAATAAC 663
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 797 CCGACTGACACCCGACACGCCCTACCTCAANTGTGGATTATGTATCATGAACCTACTGGAAG 856
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 664 CCTACTGAGACTGGTAGCAATGCCTACCTCAATGTGGATTATGTATGAACCTATTGGAAG 733
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 857 GACAATGGAGCACCCAGCTGAGAACTCATCGTTGGATTCCCTACCTATGGACACACAACCTTC 916
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 724 AACATGGAGCCCAGCTGAGAGCTCATTTGGATTCCAGAGITGGACACACCTTC 763
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 917 ATCCTGAGCAACCCCTCCAACACTGGAAATGGTGCCCCCAC 957
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 784 ATCCTGAGAAACCCTCTGATAATGGATTGGTGGTCCCCCTAC 824
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 11

B1518221/c

LOCUS 603042049t1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5182357 3', mRNA sequence.

DEFINITION B1518221

ACCESSION B1518221

VERSION B1518221.1 GI:15343013

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 824)

AUTHORS NIH-MGC http://mgc.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)

JOURNAL Contact: Robert Strausberg, Ph.D.

COMMENT Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM1455 row: a column: 14
High quality sequence start: 24
High quality sequence stop: 746.

FEATURES

Location/Qualifiers

1..824

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5182357"

/clone.lib="NIH_MGC_116"

/lab_host="DH10B"

/note="Organ: pooled colon, kidney, stomach; Vector: pCMV-SPORT6; Site:1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 3 colons, age 26 yo male, 49 yo female, 71 yo male colon; 46 yo male kidney, and pool of 2 stomachs, 62 yo male and 70 yo female. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 023. Note: this is a NIH_MGC Library."

BASE COUNT 176 a 228 c 227 g 193 t

ORIGIN

Query Match 39.5%; Score 641.2; DB 13; Length 824;
Best Local Similarity 95.9%; Pred. No. 1.6e-167;
Matches 733; Conservative 0; Mismatches 23; Indels 8; Gaps 7;

Ov 819 CCTACCTCAATGTGGATTATGTATGCATGAACCT-ACCTGGAAGGACAATGGAGCACCAGCTGAG 877


```
QY 824 CTCAATGTGG-ATTATGTCATGAACATA-CTGGAAGGACAAATGAGCAGCAGCTGAGAAGC 881
|||||
Db 721 CTCAATGTGGCATATGTCATGAACATATCTGGAAGAACAAATGAGGCCAGCTGAGAAGC 780
|||||
QY 882 TCATCGTGTGGATTCCTTACCTATGGAACACACATTCATCTCGACACACCC 931
|||||
Db 781 TCATGTGTGGATTCCTCGAGATGGAACACACCTTCATCTCGAGAAACCC 830
|||||

RESULT 13
BG868741 837 bp mRNA linear EST 29-MAY-2001
LOCUS 602787570F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4913800 5',
DEFINITION mRNA sequence.
ACCESSION BG868741
VERSION BG868741.1 GI:14219281
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 837)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapb@remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10819 row: k column: 17
High quality sequence stop: 836.
Location/Qualifiers
1. 837
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4913800"
/lab_host="NCI_CGAP_SG2"
/notes="Organ: salivary gland; Vector: pCMV-SPORT6; Site: 1;
NotI; Site: 2; SalI; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 212 a 211 c 197 g 217 t
ORIGIN

Query Match 38.7%; Score 629; DB 12; Length 837;
Best Local Similarity 84.5%; Pred. No. 4.1e-164;
Matches 707; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 102 CCATGACAAAGCTATTCCTCCACAGGTCTTGTCTTACTATGAAATTCAGCTCGGCT 161
|||||
Db 1 CGATGCCCAAGCTACTCTCTGTCACAGGTCTGGCTCTTCTGCTGAATGCTCAGCTGGGT 60
|||||
QY 162 CTGCTTACCAGCTGACATGCTACTTACCACACTGGCCAGTACCGCCAGGCTGGGC 221
|||||
Db 61 CTGCTTACATCTGATGATGCTATTTTCCAACTGGCCAGTATCGCCAGGTCTGGGA 120
|||||
QY 222 GCTTCATGCTGACAACTGACCCCTGCTCTGTACCCACCTGATCTACGCTTTGCTG 281
|||||
Db 121 GCTTCAAGCTGATGACATACCCCTGCTGTGTAATCTACCTGATCTATGCTTTCG 180
|||||
QY 282 GGAGGACGAACACGAGATACACCATCGAATGGAACGATGTGACTCTTACCAGCTT 341
|||||
Db 181 GGATGACGAACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240
|||||
QY 342 TCAATGGCTGAAATATAGACAGCCAGCTGAAACTCTCTGCGCATGTGAGGCTTGA 401
|||||
```

```
Db 241 TCAATGACTTGAAGAAACAGAGAAACAGAAACCTCTCTGGCAATTTGGAGGTGGA 300
QY 402 ACTTTCGGGACTGCCCTTTTCACTGCGCATGGTTTCTACTCTGAGAACCCGCGACACTTCA 461
|||||
Db 301 ACTTTGGAAGTCTCTCTTCACTACCATGGTTTCCACTTCTCAGAACCCGCGACACTTCA 360
|||||
QY 462 TCACCTCAGTCATCAAAATTCCTCGCCAGTATGAGTTTGACGGCTGGAGCTTTGACTGGG 521
|||||
Db 361 TTACCTCAGTCATCAAAATTCCTCGCTCAGTATGGTTTGATGGACTGGACCTGAGTGGG 420
|||||
QY 522 AGTACCTCGCTCTCGTGGGAGCCCTCTCCTCAGGACAAAGCATCTCTTCACTGTCTGGTGC 581
|||||
Db 421 NATACCCAGGCTCAGCTGGGAGCCCTCTCCTCAGGACAAAGCATCTCTTCACTGTCTGGTGA 480
|||||
QY 582 AGGAAATGCTGAAGCTTTTGACGAGGAGCCCAAGCAGATCAACAAGCCCAAGCTGATGG 641
|||||
Db 481 AGGAAATGCTGAAGCTTTTGACGAGGAGGCTATTGAGACCAACACAGCCCAAGCTGATGG 540
|||||
QY 642 TCACCTGCTGCAGTGTGCTGGCATCTCCAATATCCAGTCTTGCTATGAGATCCCCCAAC 701
|||||
Db 541 TTACTGCTGCTGTAGCTGGTGGATTTCCAACATCCAGGCTGGCTATGAGATCCCTGAAC 600
|||||
QY 702 TGTACAGTACCTGGACTACATCCATGTCATGACCTACGACCTCCATGCTCTCTGGAGG 761
|||||
Db 601 TTTCTAAGTACCTGGATTTTCATCATGTATGATGATGATGATGATGATGATGATGATG 660
|||||
QY 762 GCTACACTGGAGAGAACAGCCCTCTACAAATACCCGACTGACACCGGACGCAAGCCCT 821
|||||
Db 661 GCTACACTGGGAGATAGTCTCTTTACAAATACCCCTACTGAGACTGGTATGAGATGCCT 720
|||||
QY 822 ACCTCAATGTGGATTATGTCATGAACTACTGGAAGGACAAATGAGGACCAAGCTGAGAAGC 881
|||||
Db 721 ACCTCAATGTGGATTATGTCATGAACTATTGGAAGAACAAATGAGGACCAAGCTGAGAAGC 780
|||||
QY 882 TCATCTGCTGGATTCCTTACCTATGAGACAACTTCTCTGACACACCTCTCCCAAC 938
|||||
Db 781 TCATTTGTTGATTTCCAGAGATGAGACACCTTCAATCTCTGAGAAACCTCTTGATA 837
|||||
```

RESULT 14
BG866775

LOCUS 602786605F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4912608 5',
DEFINITION mRNA sequence.

ACCESSION BG866775

VERSION BG866775.1 GI:14217315

KEYWORDS EST.

SOURCE house mouse.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 847)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapb@remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM10816 row: j column: 01

High quality sequence stop: 774.

FEATURES

source

1. 847

Location/Qualifiers

/organism="Mus musculus"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="IMAGE:4912608"

/clone_lib="NCI_CGAP_SG2"

/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site:1;
Not1; Site:2; SalI; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI-CGAP Library."
BASE COUNT 214 a 211 c 201 g 221 t
ORIGIN

Query Match 38.5%; Score 626; DB 12; Length 847;
Best Local Similarity 84.5%; Pred. No. 2.8e-163; Indels 1; Gaps 1;
Matches 715; Conservative 0; Mismatches 130;
Qy 100 AACCATGACAAAGCTATTCTCCACAGGCTTGTCTTATCTAGTGAATTCAGCTCGG 159
Db 2 AACGATGCCAAGCTACTCTCTCGACAGGCTGGCTCTCTGCTGAATGCTCAGCTGG 61
Qy 160 CTCGTCCATGCTGACATCTACTTACCACTGGGCCAGTACCGGCCAGGCTGGG 219
Db 62 GTCTGCCATCAATCTGATATGCTATTTCCCACTGGGCCAGTATCGCCAGGCTGGG 121
Qy 220 GCGTTCATGCTGACAACTCGACCCCTGCTGTACCCACCTGATCTAGCTTGGC 279
Db 122 GAGTTTCAGCTGATGATGATTAACCCCTGCTGTACTCACCTGATCTATGCTTTGC 181
Qy 280 TGGGAGGAGAACACGAGATCAGCACCATCGAATGGAACGATGTGACTCTCTACCAAGC 339
Db 182 TGGGATGAGAACATGAGATCACCACCATAGATGAATGATGTTACTCTCTATAAGC 241
Qy 340 TTTCAATGGCTGAAATTAAGACAGACGACCTGAAACTCTCTCTGGCCATGGAGGCTG 399
Db 242 TTTCAATGACTTGAATTAAGACAGACGACCTGAAACTCTCTCTGGCCATGGAGGCTG 301
Qy 400 GAATTCGGGACTGCCCTTCTCACTGCTGCTGCTTCTACTCTCTGAGACCCGACGCTTT 459
Db 302 GAATTCGGGACTGCCCTTCTCACTGCTGCTGCTTCTACTCTCTGAGACCCGACGCTTT 361
Qy 460 CATCACTCAGTCAATCAATTCCTGGCCAGTATGAGTTGACGGCTGGAGCTTTGACTG 519
Db 362 CATCACTCAGTCAATCAATTCCTGGCCAGTATGAGTTGACGGCTGGAGCTTTGACTG 421
Qy 520 GGAGTACCTGGCTCTGGGGAGCCCTCTCTCAGGACGAGTATCTTCACTGCTCTGGT 579
Db 422 GGAATACCCAGGCTCAGCTGGGAGCCCTCTCTCAGGACGAGTATCTTCACTGCTCTGGT 481
Qy 580 GCAGGAATGCTGAAGCTTTTGAGCAGGAGGCAAGCAGATCAACAGCCAGGCTGAT 639
Db 482 GAAGGAATGCTGAAGCTTTTGAGCAGGAGGCTTTGAGAGACACAGCCGACGCTGAT 541
Qy 640 GGTCACTGCTGAGTGTGCTGGCATCTCCAAATATCCAGTCTGGCTATGAGATCCCCCA 699
Db 542 GGTCACTGCTGAGTGTGCTGGCATCTCCAAATATCCAGTCTGGCTATGAGATCCCCCA 601
Qy 700 ACTGTCAGTACCTGGACTACATCTGATGATGATGATGATGATGATGATGATGATGAT 759
Db 602 ACTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 661
Qy 760 GGGTACACTGGAGAGAACCCCTCTCAAAATACCCGACTGACACCGCAGCAGCAGC 819
Db 662 GGGTACACTGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 721
Qy 820 CTACTCAATGTGGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 879
Db 722 CTACTCAATGTGGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
Qy 880 GCTCATCTGTTGATTCCTTACTATGAGACAACTTCACTCTGAGCAACCCCTGCAACAGC 939
Db 781 GCTCATCTGTTGATTCCTTACTATGAGACAACTTCACTCTGAGCAACCCCTGCAACAGC 840
Qy 940 TGGAT 945
Db 841 GGAAT 846

RESULT 15
BG866417
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
BASE COUNT
ORIGIN
Query Match 38.4%; Score 623.2; DB 12; Length 809;
Best Local Similarity 86.4%; Pred. No. 1.7e-162;
Matches 699; Conservative 0; Mismatches 109; Indels 1; Gaps 1;
Qy 376 AACTCTCTGCGCATTTGGAGGCTGGAACCTTGGGACTGCCCCCTTTCAGTGCATGTTTC 435
Db 1 AACCTCTCTGCAATTTGGAGGCTGGAACCTTGGGACTGCCCCCTTTCAGTGCATGTTTC 60
Qy 436 TACTCTCTGAGAACGCCAGACTTTTCATCACCTCAGTCAATCAATTTCTCGGCGCAGTATGA 495
Db 61 CACTTCTCAGAACCCCGACGCTTTCATTCCTCAGTCAATCAATTTCTCGGCGCAGTATGG 120
Qy 496 GTTTGACGGCTGGAGCTTTGACTGGGAGTACCTGCTCTGCTGGGAGCCCTCTCTCAGGA 555
Db 121 GTTTGATGAGTGGAGCTGGAGTACCCAGGCTCAGCTGGGAGCCCTCTCTCAGGA 180
Qy 556 CAAGCATCTTTCAGTCTGCTGGTGCAGGAATGGCTGAAGCTTTTGGAGCAGGAGCCAA 615
Db 181 CAAGCATCTTTCAGTCTGCTGGTGCAGGAATGGCTGAAGCTTTTGGAGCAGGAGGCTAT 240
Qy 616 GCAGATCAACAGCCCGAGCTGATGCTCAGTCTGCTCAGTCTGCTGCTGCTGCTGCTGCTGCT 675
Db 241 TGAGAGCAACAGCCCGAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
Qy 676 CCAGTCTGGCTATGAGATCCCGCAACTGCTCAGAGTACCTGGAGTACATCCATGCTGATGAC 735
Db 301 CCAGCTGGCTATGAGATCCCGCAACTGCTCAGAGTACCTGGAGTACATCCATGCTGATGAC 360
Qy 736 CTAGACCTCCATGCTCTGGGAGGCTTACACTGGAGAGACACGCCCTCTCTACAAATA 795
Db 361 ATATGACCTCCATGCTCTGGGAGGCTTACACTGGGAGAGTACTCTCTCTTACAAATA 420

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2003, 02:31:01 ; Search time 4011.96 Seconds
(without alignments)
11062.367 Million cell updates/sec

Title: US-10-004-219b-3
Perfect score: 1525
Sequence: 1 atggcagactactctctgt.....taaaattgttagcacaaca 1525

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg_hum.*
- 31: em.htg_inv.*
- 32: em.htg_other.*
- 33: em.htg_mus.*
- 34: em.htg_pln.*
- 35: em.htg_rod.*
- 36: em.htg_mam.*
- 37: em.htg_vrt.*
- 38: em.sy.*
- 39: em.htgo_hum.*
- 40: em.htgo_mus.*
- 41: em.htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1525	100.0	1530	10	AF290003	AF290003 Mus muscu
2	1520.2	99.7	1557	10	BC034548	BC034548 Mus muscu
3	1519.8	99.7	1538	10	BC011134	BC011134 Mus muscu
4	1517.2	99.5	1526	10	AF154571	AF154571 Mus muscu
5	1041	68.3	1625	9	AF290004	AF290004 Homo sapi
6	959.2	62.9	1529	4	AB051629	AB051629 Bos tauru
7	856	56.1	1354	9	AB025008	AB025008 Homo sapi
8	854	56.0	1369	6	AX405989	AX405989 Sequence
9	714.6	46.9	1525	10	M34584	M34584 Mus musculu
10	713	46.8	1506	10	D87757	D87757 Mus musculu
11	708.2	46.4	1209	10	AY049765	AY049765 Mus muscu
12	706.6	46.3	1209	10	AY065557	AY065557 Mus muscu
13	669	43.9	4250	9	AK098814	AK098814 Homo sapi
14	665	43.6	1188	9	AB025009	AB025009 Homo sapi
15	460	30.2	994	10	MMU58900	U58900 Mus musculu
16	430.4	28.2	1562	5	BJA345054	AJ345054 Bufo japo
17	426.8	28.0	1633	9	HSU29615	U29615 Human chito
18	426.8	28.0	1636	6	AR206041	AR206041 Sequence
19	426.8	28.0	1636	6	AR212138	AR212138 Sequence
20	426.8	28.0	1636	6	AX108750	AX108750 Sequence
21	426.8	28.0	1643	6	AR172107	AR172107 Sequence
22	426.8	28.0	1710	9	HSU62862	U62862 Homo sapien
23	426.8	28.0	1713	6	AR172108	AR172108 Sequence
24	425.2	27.9	1656	6	AR206042	AR206042 Sequence
25	425.2	27.9	1656	6	AR212139	AR212139 Sequence
26	425.2	27.9	1656	6	AX108752	AX108752 Sequence
27	371.4	24.4	2237	9	MMU87259	U87259 Macaca mula
28	368.2	24.1	2198	9	HSU09550	U09550 Human ovidu
29	365.2	23.9	2525	10	MDSOGP	D32137 Mouse mRNa
30	365	23.9	1406	9	AF484550	AF484550 Macaca ra
31	363.4	23.8	2228	9	BABEDDSG	M59903 Papio hamad
32	362	23.7	2504	6	E09046	E09046 cDNA encodi
33	355.6	23.3	2353	10	MAU15048	U15048 Mesocricetu
34	355.6	23.3	2387	10	HAMOGP	D32218 Hamster mRN
35	350.8	23.0	2366	6	E09047	E09047 cDNA encodi
36	350.4	23.0	233175	2	AC125461	AC125461 Mus muscu
37	343.2	22.5	1433	6	AR042834	AR042834 Sequence
38	343.2	22.5	1433	9	HSU58514	U58514 Human chiti
39	341.6	22.4	1418	9	HSU49825	U49835 Human YKL-3
40	341.6	22.4	1449	9	BC011460	BC011460 Homo sapi
41	340.4	22.3	1500	9	HSU58515	U58515 Human chiti
42	339.4	22.3	1994	4	BOV95BOGP	D16639 Bovine mRNa
43	339.4	22.3	1994	6	E09045	E09045 cDNA encodi
44	333.2	21.8	1599	9	AK055165	AK055165 Homo sapi
45	331.4	21.7	2034	4	OAU16719	U16719 Ovis aries

ALIGNMENTS

RESULT 1
AF290003
LOCUS Mus musculus acidic mammalian Chitinase precursor, mRNA, complete cds.
DEFINITION AF290003
ACCESSION AF290003
VERSION AF290003.1 GI:12597290
KEYWORDS
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 1530)
AUTHORS Boot,R.G., Blommaert,E.F., Swart,E., Ghauharali-van der Vlugt,K., Bijl,N., Moe,C., Place,A. and Aerts,J.M.

AF290003 1530 bp mRNA linear ROD 27-FEB-2001
Mus musculus acidic mammalian Chitinase precursor, mRNA, complete cds.

TITLE Identification of a novel acidic mammalian chitinase distinct from chitotriosidase
J. Biol. Chem. 276 (9), 6770-6778 (2001) *Man*
MEDLINE 21125893
PUBMED 11085997
REFERENCE 2 (bases 1 to 1530)
AUTHORS Boot, R.G., Verhoeck, M., Swart, E. and Aerts, J.M.
TITLE Direct Submission
JOURNAL Submitted (26-JUL-2000) Dept. of Biochemistry, Academic Medical Center, University of Amsterdam, Meibergdreef 15, Amsterdam 1105 AZ, The Netherlands
FEATURES
source Location/Qualifiers
1. .1530
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
/sex="male"
/tissue_type="lung"
/dev_stage="9-11 weeks"
3. .1424
/EC_number="3.2.1.14"
/note="chitinase precursor; AMCase"
/codon_start=1
/product="acidic mammalian chitinase precursor"
/protein_id="AAG60018.1"
/db_xref="GI:12597291"
/translation="MAKLIVTGLALLNLAQLGSAYNLICYFTNWAQYRPLGSPKPD
DINPCLCTHLIVAFAGMONNEITIEWVDVLYKAPNDLKNRSKILTLAIGWNGF
TAPFTWVTSQNRQFISVIFKIQYGFDDGLDWEIPGSRSPDKHLFVLVK
EMREAEQEAIESNPRFLMVAAGISINQAGYEIPELSKYLDLTHVMYDHLGSM
EGYTGNSPLKYPTGTGSNAYLNVDYVMYWNKNAPAEKLIIVGPEYGHTEILRNP
SDNGIGAPTSDDGPAGAYTRQAGWAYEICTFLRSRGATEVMDSEVPYAYKANLW
GYDNIKSFVKAOWLKONNFGAMIAWIDLDFTGFCDOGKFLPLSTLNKALGISTE
GCTAPDVPSPEPTTPGSGSGSGSGSGSGSGGFCADKADGLYPVADDRNAPQCING
ITVQHQAGLVFDTSCNENWF"
BASE COUNT 382 a 383 c 366 g 399 t
ORIGIN
Query Match 100.0%; Score 1525; DB 10; Length 1530;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1525; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGGCCAGTACTTCTGTCACAGGTCGGCTTCTTCTGCTGAATGCTCAGCTGGGTCT 60
Db |||||||
Qy 3 ATGGCCAGTACTTCTGTCACAGGTCGGCTTCTTCTGCTGAATGCTCAGCTGGGTCT 62
Db |||||||
Qy 61 GCCTACATCTGATATGCTATTTCACCACTGGGCCAGTATCGGCCAGTCTGGGAGC 120
Db |||||||
Qy 63 GCCTACATCTGATATGCTATTTCACCACTGGGCCAGTATCGGCCAGTCTGGGAGC 122
Db |||||||
Qy 121 TTCAAGCCTGATGACATTAACCCCTGCTGTACTACCTGATCTATGCTTTTGTGGG 180
Db |||||||
Qy 123 TTCAAGCCTGATGACATTAACCCCTGCTGTACTACCTGATCTATGCTTTGTGGG 182
Db |||||||
Qy 181 ATGACAGCAATGAGATCACCACATAGAAATGAATGATGTTACTCTCTATAAGCTTTC 240
Db |||||||
Qy 183 ATGACAGCAATGAGATCACCACATAGAAATGAATGATGTTACTCTCTATAAGCTTTC 242
Db |||||||
Qy 241 AATGACTTGAACAAAGGACAGCAACAACTGAAACCCCTCTGCAATTTGGAGGCTGAAC 300
Db |||||||
Qy 243 AATGACTTGAACAAAGGACAGCAACAACTGAAACCCCTCTGCAATTTGGAGGCTGAAC 302
Db |||||||
Qy 301 TTTTGAAGTCTCTCTTCACTACCATGTTTCCACTTCTCAGAACCGCCAGACCTTCATT 360
Db |||||||
Qy 303 TTTTGAAGTCTCTCTTCACTACCATGTTTCCACTTCTCAGAACCGCCAGACCTTCATT 362
Db |||||||
Qy 361 ACCTCAGTCTCAAAATTCGCTCAGTATGGGTTTGATGGAGTGGACCTGGAGCTGGAA 420
Db |||||||
Qy 363 ACCTCAGTCTCAAAATTCGCTCAGTATGGGTTTGATGGAGTGGACCTGGAGCTGGAA 422
Db |||||||
Qy 421 TACCCAGGCTCAGTGGAGCCCTCTCTCAGGACAAGCATCTTCTACTGTCCTGGTGAAG 480
Db |||||||
Qy 423 TACCCAGGCTCAGTGGAGCCCTCTCTCAGGACAAGCATCTTCTACTGTCCTGGGGAAG 482
Db |||||||

Qy 481 GAAATGCGTGAAGCTTTTGAGCAGGAGGCTATTGAGAGCAACAGGCCAGCTGATGTT 540
Db |||||||
Qy 483 GAATCGGTGAAGCTTTTGAGCAGGAGGCTATTGAGAGCAACAGGCCAGCTGATGTT 542
Db |||||||
Qy 541 ACTGCTGCTAGCTGGTGGGATTTCCAAACATCAGGCTGGCTATGAGATCCCTGAACCT 600
Db |||||||
Qy 543 ACTGCTGCTAGCTGGTGGGATTTCCAAACATCAGGCTGGCTATGAGATCCCTGAACCT 602
Db |||||||
Qy 601 TCTAAGTACCTGATTTTCATCCATGTCATGACATATGACCTCCATGGCTCTCTGGAGGC 660
Db |||||||
Qy 603 TCTAAGTACCTGATTTTCATCCATGTCATGACATATGACCTCCATGGCTCTCTGGAGGC 662
Db |||||||
Qy 661 TACATGGGGAGAAATAGTCTCTTTACAAATACCTACTCAGACTGGTAGCAATGCTAC 720
Db |||||||
Qy 663 TACATGGGGAGAAATAGTCTCTTTACAAATACCTACTCAGACTGGTAGCAATGCTAC 722
Db |||||||
Qy 721 CTCATGTGGATTTATGTCATGAATATTTGGAAGAAATGAGCCCGAGCTGAGAGCTC 780
Db |||||||
Qy 723 CTCATGTGGATTTATGTCATGAATATTTGGAAGAAATGAGCCCGAGCTGAGAGCTC 782
Db |||||||
Qy 781 ATTGTTGGATTTCCAGAGTATGACACACCTTCTCTGAGAACCCCTCTGATATGGA 840
Db |||||||
Qy 783 ATTGTTGGATTTCCAGAGTATGACACACCTTCTCTGAGAACCCCTCTGATATGGA 842
Db |||||||
Qy 841 ATTGTTGGATTTCCAGAGTATGACACACCTTCTCTGAGAACCCCTCTGATATGGA 900
Db |||||||
Qy 843 ATTGTTGGATTTCCAGAGTATGACACACCTTCTCTGAGAACCCCTCTGATATGGA 902
Db |||||||
Qy 901 TGGGCTCTACTATGAGATTTGACACCTTTCTGAGAAAGTGGAGCCACTGAGGTCTGGAGTC 960
Db |||||||
Qy 903 TGGGCTCTACTATGAGATTTGACACCTTTCTGAGAAAGTGGAGCCACTGAGGTCTGGAGTC 962
Db |||||||
Qy 961 TCCCAAGAGTGGCTCTGCTATAGGCAACAGTGGCTTGGCTATGACAAATATCAAG 1020
Db |||||||
Qy 963 TCCCAAGAGTGGCTCTGCTATAGGCAACAGTGGCTTGGCTATGACAAATATCAAG 1022
Db |||||||
Qy 1021 AGCTTCAAGTGTAAAGGCTCAGTGGCTTAAAGCAACAAATTTTGGAGTGGCATGATG 1080
Db |||||||
Qy 1023 AGCTTCAAGTGTAAAGGCTCAGTGGCTTAAAGCAACAAATTTTGGAGTGGCATGATG 1082
Db |||||||
Qy 1081 GCATTTGACCTTGTATGACTTCACTGGCTCTTCTGATGATGAGGAAATTTCTCTGACT 1140
Db |||||||
Qy 1083 GCATTTGACCTTGTATGACTTCACTGGCTCTTCTGATGATGAGGAAATTTCTCTGACT 1142
Db |||||||
Qy 1141 TCTACTTTGAAACAAAGCCCTTGGCATATCCACTGAAAGTTGACAGCTCTCTGAGTGCCT 1200
Db |||||||
Qy 1143 TCTACTTTGAAACAAAGCCCTTGGCATATCCACTGAAAGTTGACAGCTCTCTGAGTGCCT 1202
Db |||||||
Qy 1201 TCCGAGCCAGTACTTCTCTCAGGAAAGTGGAGTGGGGTGGAAAGCTCCGGAGGAAGC 1260
Db |||||||
Qy 1203 TCCGAGCCAGTACTTCTCTCAGGAAAGTGGAGTGGGGTGGAAAGCTCCGGAGGAAGC 1262
Db |||||||
Qy 1261 TCTGAGGAGTGGATTTCTGTCGAGCAAGAGAGATGGCTCTTACCTGTGGCAGATGAC 1320
Db |||||||
Qy 1263 TCTGAGGAGTGGATTTCTGTCGAGCAAGAGAGATGGCTCTTACCTGTGGCAGATGAC 1322
Db |||||||
Qy 1321 AGAAATGCTTTTGGCAGTGCATCAATGGAATCACATACCAGCAGCATTTGCAAGCAGGG 1380
Db |||||||
Qy 1323 AGAAATGCTTTTGGCAGTGCATCAATGGAATCACATACCAGCAGCATTTGCAAGCAGGG 1382
Db |||||||
Qy 1381 CTTGTTTTGATACCAAGCTGTAATGCTGCACTGGCCATGAAACCTTAATGCCATTTCTTC 1440
Db |||||||
Qy 1383 CTTGTTTTGATACCAAGCTGTAATGCTGCACTGGCCATGAAACCTTAATGCCATTTCTTC 1442
Db |||||||
Qy 1441 AGAAATTTCTGCACTCTCTTTTACTCTTACCCAAAGAGTAACTATCTTCCCTTTAACCTTA 1500
Db |||||||
Qy 1443 AGAAATTTCTGCACTCTCTTTTACTCTTACCCAAAGAGTAACTATCTTCCCTTTAACCTTA 1502
Db |||||||
Qy 1501 TGAATAAATAATGGTAGCCAAAACA 1525
Db |||||||
Qy 1503 TGAATAAATAATGGTAGCCAAAACA 1527
Db |||||||


```
1084 GCCATTGACCTTGATGACTTCACTGGCTCTTCTGTGATCAGGAAAAATTCCTCTGACT 1143
1141 TCTACTTTGAACAAAGCCCTTGGCATATCCACTGAAGTTGACAGCTCTCTGACGTCCT 1200
1144 TCTACTTTGAACAAAGCCCTTGGCATATCCACTGAAGTTGACAGCTCTCTGACGTCCT 1203
1201 TCCGAGCAGTACTCTCTCCAGGAAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGG 1260
1204 TCCGAGCAGTACTCTCTCCAGGAAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGG 1263
1261 TCTGGAGCAGTACTCTCTCCAGGAAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGG 1320
1264 TCTGGAGCAGTACTCTCTCCAGGAAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGG 1323
1321 AGAATGCTTTTGGGAGTGCATCAATGGAATCACATACAGCAGGAGTTCGACAGGAGG 1380
1324 AGAATGCTTTTGGGAGTGCATCAATGGAATCACATACAGCAGGAGTTCGACAGGAGG 1383
1381 CTGTTTTTTCATACAGTCTTAATGCTGCAACTGGCCATGAACCTTAATGCCATTTCTTC 1440
1384 CTGTTTTTTCATACAGTCTTAATGCTGCAACTGGCCATGAACCTTAATGCCATTTCTTC 1443
1441 AGAATTTTCGACTCTCTCTTACTCTCTCACCAGGAGTAACTATCTCTCTTAACTTA 1500
1444 AGAATTTTCGACTCTCTCTTACTCTCTCACCAGGAGTAACTATCTCTCTTAACTTA 1503
1501 TGCATAAATAATGAGTGCACCAAAACA 1525
1504 TGCATAAATAATGAGTGCACCAAAACA 1528
```

RESULT 3

```
BC011134 1538 bp mRNA linear ROD 07-AUG-2002
LOCUS MGC:18771 IMAGE:4165150, mRNA, complete cds.
DEFINITION Mus musculus, similar to eosinophil chemotactic cytokine, clone
BC011134
VERSION BC011134.1 GI:15029821
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1538)
TITLE Strausberg, R.
JOURNAL Direct Submission
Submitted (25-JUL-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amgbcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK plate: 24 Row: e Column: 11.
FEATURES
source
1..1538
/organism="Mus musculus"
/db_xref="taxon:10090"
```

```
/map="FVB/N"
/clone="MGC:18771 IMAGE:4165150"
/tissue_type="Salivary gland, 10 week old female mouse"
/clone_lib="NCI_CGAP_SG2"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
323..1420
CDS
/codon_start=1
/product="Similar to eosinophil chemotactic cytokine"
/protein_id="AAH1134.1"
/db_xref="GI:15029822"
HFTVLVKEMREAFQEAEIENSRPLMTVAAGVINSIQAGYEIPSLKLYDFIHM
translation="MSTSONRQTFITSVIKFLROYGFDGLDLWEYPSGRSGPPQDK
HYFLILRNPDNDGIGAPTSQDGPAGYTRQAGWAYEICTFLBSGATEWDAQSEVPY
AYKANWLGIDNKSFSVRAKWLQNNFGAMIAIDLDDFTGSPGCDQKFPITSLN
KAUGISTEGTADVPSEPVTPPPGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSG
AFWQCINGITYQQHQAQGLVFTSCNCCNP"
BASE COUNT 393 a 382 c 365 g 398 t
ORIGIN
Query Match 99.7%; Score 1519.8; DB 10; Length 1538;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1521; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 3 GGCAAGCTACTTCTCGTCACAGGCTCGGCTCTTCTGTGTAATGCTCAAGTCTGAGTGGGCTGTCG 62
DB 1 GGCAAGCTACTTCTCGTCACAGGCTCGGCTCTTCTGTGTAATGCTCAAGTCTGAGTGGGCTGTCG 60
QY 63 CTACAATCTGATATGCTATTTCACCAACTGGGCGCCAGTAGTCGCCAGGCTCGGAGCTT 122
DB 61 CTACAATCTGATATGCTATTTCACCAACTGGGCGCCAGTAGTCGCCAGGCTCGGAGCTT 120
QY 123 CAAGCTGATGACATTAACCCCTGCTGTCTACTACCTACCTGATCTATGCTTCTGCTGGAT 182
DB 121 CAAGCTGATGACATTAACCCCTGCTGTCTACTACCTGATCTATGCTTCTGCTGGAT 180
QY 183 GCAGAACATGAGATCACCACCATAGAAATGGAATGATGTTACTCTCTATAAAGCTTTCAA 242
DB 181 GCAGAACATGAGATCACCACCATAGAAATGGAATGATGTTACTCTCTATAAAGCTTTCAA 240
QY 243 TGACTTGAACAAACAGGACAGCAAACTGAAACCCCTCTCGGCAATTTGGAGGCTTGAACCTT 302
DB 241 TGACTTGAACAAACAGGACAGCAAACTGAAACCCCTCTCGGCAATTTGGAGGCTTGAACCTT 300
QY 303 TGGAACTGCTCTCTTCTCACTACCATGTTTCCACTTCTCAGAACCCGACAGCTTCAATAC 362
DB 301 TGGAACTGCTCTCTTCTCACTACCATGTTTCCACTTCTCAGAACCCGACAGCTTCAATAC 360
QY 363 CTCAGTCATCAAAATTTCTCGTCAGTATGGGTTTGTGAGTGGAGCTGGAGCTGGGAATA 422
DB 361 CTCAGTCATCAAAATTTCTCGTCAGTATGGGTTTGTGAGTGGAGCTGGAGCTGGGAATA 420
QY 423 CCCAGGCTCAGTGGGAGGCTCTCTCAGGACAGCATCTCTCTACCTGCTCTGCTGTAAGGA 482
DB 421 CCCAGGCTCAGTGGGAGGCTCTCTCAGGACAGCATCTCTCTACCTGCTCTGCTGTAAGGA 480
QY 483 AATGCTGTAAGCTTTTGGACAGGAGCTATTGAGAGCAACAGCCGACAGCTGATGCTTAC 542
DB 481 AATGCTGTAAGCTTTTGGACAGGAGCTATTGAGAGCAACAGCCGACAGCTGATGCTTAC 540
QY 543 TGTGCTGTAGCTTGGTGGGATTTCCAACTCCAGGCTGGCTATGAGATCCCTGAACCTTTC 602
DB 541 TGTGCTGTAGCTTGGTGGGATTTCCAACTCCAGGCTGGCTATGAGATCCCTGAACCTTTC 600
QY 603 TAAGTACCTGGATTTTCATCATGTCATGATGATGATGATGATGATGATGATGATGATGATGAT 662
DB 601 TAAGTACCTGGATTTTCATCATGTCATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 663 CACTGGGAGAGATAGTCTCTTTTACAAATACCTCTACTGAGAGCTGGTAGCAATGCCACCT 722
DB 661 CACTGGGAGAGATAGTCTCTTTTACAAATACCTCTACTGAGAGCTGGTAGCAATGCCACCT 720
```


Db	361	TCAGTCATCAAAATTTCTGCGTCAGTATGGTTTGATGGAGCTGGACCTGGGGAATAC	420
QY	424	CCAGGCTCAGTGGGAGCCCTCTCAGGACACAGCATCTCTTCACTGCTGCTGGTGAAGGAA	483
Db	421	CCAGGCTCAGTGGGAGCCCTCTCAGGACACAGCATCTCTTCACTGCTGCTGGTGAAGGAA	480
QY	484	ATGCGTGAAGCTTTTTCAGCAGGAGGCTATTGAGACCAAGAGCCAGAGCTGATGGTTACT	543
Db	481	ATGCGTGAAGCTTTTTCAGCAGGAGGCTATTGAGACCAAGAGCCAGAGCTGATGGTTACT	540
QY	544	GCTGCTGATGCTGGGATTTTCAACATCCAGGCTGGGTATGAGATCCCTGAACCTTCT	603
Db	541	GCTGCTGATGCTGGGATTTTCAACATCCAGGCTGGGTATGAGATCCCTGAACCTTCT	600
QY	604	AAGTACCTGGATTCATCCATGATGACATATGACCTCCATGGCTCTGGGAGGCTAC	663
Db	601	AAGTACCTGGATTCATCCATGATGACATATGACCTCCATGGCTCTGGGAGGCTAC	660
QY	664	ACTGGGAGATAGTCTCTTTTACAAATACCTTACTGAGACTGGTAGCAATGCTACCTC	723
Db	661	ACTGGGAGATAGTCTCTTTTACAAATACCTTACTGAGACTGGTAGCAATGCTACCTC	720
QY	724	ATGCGGATATGCTATGAACTATTGGAAGACAAATGGAGCCCGAGCTGAGAGCTCAT	783
Db	721	AATGTGGATATGCTATGAACTATTGGAAGACAAATGGAGCCCGAGCTGAGAGCTCAT	780
QY	784	GTTGGATTCAGAGATGAGACACCTTCATCTGAGAACCCCTCTCATATGAAAT	843
Db	781	GTTGGATTCAGAGATGAGACACCTTCATCTGAGAACCCCTCTCATATGAAAT	840
QY	844	GTTGCCCTTACCTCTGGTGTGATGCCCTGCTGGCCCTATACACAGACAGCTGGTCTGG	903
Db	841	GTTGCCCTTACCTCTGGTGTGATGCCCTGCTGGCCCTATACACAGACAGCTGGTCTGG	900
QY	904	GCCTACTATGAGATTTGCACCTTTCTGAGAAAGTGAGCCACAGCTGCGGATGCTCTCC	963
Db	901	GCCTACTATGAGATTTGCACCTTTCTGAGAAAGTGAGCCACAGCTGCGGATGCTCTCC	960
QY	964	CAAGAAGTCCCTATGCTTAAAGCCAAAGAGTGGCTTGGCTATGACAAATATCAAGAGC	1023
Db	961	CAAGAAGTCCCTATGCTTAAAGCCAAAGAGTGGCTTGGCTATGACAAATATCAAGAGC	1020
QY	1024	TTCAGTGTAAAGCTCAGTGGCTTAAAGCAGAACAAATTTTGGAGTGGCCATGATCTGGCC	1083
Db	1021	TTCAGTGTAAAGCTCAGTGGCTTAAAGCAGAACAAATTTTGGAGTGGCCATGATCTGGCC	1080
QY	1084	ATTGACCTTGATGACTTCACTGGCTTTCTGATGATCAGGGAATTTTCTGCTGACTTCT	1143
Db	1081	ATTGACCTTGATGACTTCACTGGCTTTCTGATGATCAGGGAATTTTCTGCTGACTTCT	1140
QY	1144	ACTTTGAACAAAGCCCTTGGCATATCCACTGAAGTTTGACAGCTCCTGACGTGCTCTCC	1203
Db	1141	ACTTTGAACAAAGCCCTTGGCATATCCACTGAAGTTTGACAGCTCCTGACGTGCTCTCC	1200
QY	1204	GAGCCAGTGACTCTCTCCAGGAAAGTGGAGTGGGGTGGAGCTCCGAGGAAGCTCT	1263
Db	1201	GAGCCAGTGACTCTCTCCAGGAAAGTGGAGTGGGGTGGAGCTCCGAGGAAGCTCT	1260
QY	1264	GGAGGAGTGGATTTCTGCCCCACAAAGCAGATGGCTTACCCTGTGGCAGATGACAGA	1323
Db	1261	GGAGGAGTGGATTTCTGCCCCACAAAGCAGATGGCTTACCCTGTGGCAGATGACAGA	1320
QY	1324	AATGCTTTTGGCAGTGATCAATGAATACATACAGCAGCATTTGTCAAGCAGGCTT	1383
Db	1321	AATGCTTTTGGCAGTGATCAATGAATACATACAGCAGCATTTGTCAAGCAGGCTT	1380
QY	1384	GTTTTTGATACAGCTGTAATTTGCTGCACTGGCCATGAACCTTAATGCTTCTCCAGA	1443
Db	1381	GTTTTTGATACAGCTGTAATTTGCTGCACTGGCCATGAACCTTAATGCTTCTCCAGA	1440
QY	1444	AATTTCTGCATCTCTTACTCTCCCTCACCACAAAGTAACCTATCTCCCTTAACTTATGC	1503
Db	1441	AATTTCTGCATCTCTTACTCTCCCTCACCACAAAGTAACCTATCTCCCTTAACTTATGC	1500
QY	1504	AATAAATTTGGTAGCCAAAACA	1525
Db	1501	AATAAATTTGGTAGCCAAAACA	1522
RESULT 5			
AF290004			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
CDS			
sig_peptide			
BASE COUNT			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 1198;			
Conservative			
Mismatches			
230;			
indels			
9;			
Gaps			
1;			
QY	1	ATGGCCACAGCTACTTCTCGTCACAGGTCTGGCTCTTCTGCTGCTGAATGCTGAGTGGGCT	60
Db	104	ATGACAAAGCTATTCTCTCTCACAGGTCTTGTCTCTATATGCTGCTGCTGCTGCTGCT	163
QY	61	GCCTACATCTGATGCTATTTACCACTCTGGGCGGCGGAGTATCGCCAGGCTGGGAGC	120
Db	164	GCCTACAGCTGATGCTATTTACCACTCTGGGCGGCGGAGTATCGCCAGGCTGGGCGC	223

```
QY 121 TTCAAGCCTGATGACATTAAACCCCTGCCTGTGTACTACCTGATCTATGCTCTTCTGGG 180
Db 224 TTATGCTGTGACAAACATGACCCCTGCCTCTGTACCCACCTGATCTAGCCCTTCTGGG 283
QY 181 ATGCAGAACATGAGATCACCACCATAGAAATGAATGATGTACTCTCTATPAAAGCTTTC 240
Db 284 AGGCAGAACACAGAGATCACCACCATGAAATGGAAGATGTGACTCTCTACCAAGCTTTC 343
QY 241 AATGACTTGAACAAACAGGAACACAACTGAAACCCCTCCCTGGCAATGGAGCTTGAAC 300
Db 344 AATGCCCTGAAAATPAGAACAGCAGCTGAAAACCTCTCTGGCCATGGAGCTTGAAC 403
QY 301 TTGGAAGCTGCTCTTCACTACCATGGTTTCCACTTCTCAGAACCCGACAGCTTCATT 360
Db 404 TTCGGGACTGCCCCCTTCACTGCCATGGTTTCTACTCTGAGAACCCGACAGCTTTCATC 463
QY 361 ACCTCAGTCAAAATTTCTGGTCAGTATGGTTTGTATGGACTGGACCTGGACTGGGA 420
Db 464 ACCTCAGTCAATCAAAATTTCTGGCCAGTATGATTTGACGGCTGGACTTTGACTGGGAG 523
QY 421 TACCACGCTACGCTGGGAGCCCTCTCAGGACAAAGCATCTCTTCACTGTCTCTGTGAAG 480
Db 524 TACCTGCTGCTGCTGGGAGCCCTCTCAGGACAGCATCTCTTCACTGTCTCTGTGAG 583
QY 481 GAAATGCGTGAAGCTTTTGAAGAGGAGCTATTTGAGAGCAACAGCCCGACAGCTGATGTT 540
Db 584 GAAATGCGTGAAGCTTTTGAAGAGGAGGCGCAAGCAGATCAACAAAGCCAGGCTGATGTT 643
QY 541 ACTGCTGCTGAGCTGGTGGGATTTCCAAATCCAGCTCAGGCTGGCTATGAGATCCCTGA 600
Db 644 ACTGCTGAGTATGCTGGCATCTCCAAATATCCAGTCTGGCTATGAGATCCCTCAACTG 703
QY 601 TCTAAGTACCTGGAATTCATCATCTCATGACATATGACCTCCATGGCTCTCTGGAGGSC 660
Db 704 TCACAGTACCTGAGCTACATCATCTCATGACCTAGACCTCCATGGCTCTCTGGAGGSC 763
QY 661 TACACTGGGAGAAATGAGTCTCTTTACAAATACCTACTGAGAGCTGGTAGCAATCCCTAC 720
Db 764 TACACTGGGAGAAACAGCCCTCTTACAAATACCCGACTGACACCCGAGCAACGCCCTAC 823
QY 721 CTCATGTGGATATGTCATGAATTTGGAAGACAAATGGAGCCCGCCAGCTCAGAGCTC 780
Db 824 CTCATGTGGATATGTCATGAATTTGGAAGAGCAATGGAGCCCGCCAGCTCAGAGCTC 883
QY 781 ATTTGTGATTCACAGATATGACACACCTTTCTATCTCTGAGAAACCCCTCTCATATGA 840
Db 884 ATCGTGTGATTCCTACCTATGACACACATTCATCTCTGAGCAACCCCTCCACACTGA 943
QY 841 ATTTGTGCCCCCTACCTCTGGTGATGGCCCTGCTGGCCCTATPACCAGACAGCTGGGTT 900
Db 944 ATTTGTGCCCCCACCCTCTGGTCTGCTGCTGGCCCTATGCCAAGGAGTCTGGGATC 1003
QY 901 TGGGCTTACTATGAGATTGACCTTCTGAGAACTGGAGCCACTGAGGTCTGGGATGCC 960
Db 1004 TGGGCTTACTAGGATCTGTACCTTCTGAAATATGGAGCCACTCAGGGATGGGATGCC 1063
QY 961 TCCCAAGAGTGCCTATGCTTATAAGGCCAACAGTGGCTTGGCTATGACAATATCAAG 1020
Db 1064 CCTCAGGAAGTGCCTTATGCCCTATCAGGCAATGTGTGGTGGCTATGACAACATCAAG 1123
QY 1021 AGCTTCAGTGTAAAGCTCAGTGGTCTTAAGCAGAACAAATTTGGAGGTGCCATGATCTGG 1080
Db 1124 AGCTTCGATATTAAGCTCAATGGCTTAAGCAGAACAAATTTGGAGGTGCCATGATCTGG 1183
QY 1081 GCCATTGACCTTATGACTTCACTGGCTCTTCTGTGATCAGCGGAAATTTCTCTGACT 1140
Db 1184 GCCATTGATCTGGATCTCACTGGCACTTTCTCAACAGGCAAGTTTCTCCCTTAATC 1243
QY 1141 TCTACTTTGAACAAAGCCCTTGGCATATCTCACTGAGGTGTGACAGCTCTCTCAGCTGCT 1200
Db 1244 TCCACCTTGAAGAGCCCTCGGCTCGAGAGTGCAGTTGACCGGCTCCACTCAGCCCC 1303
QY 1201 TCCGAGCCAGTGACTACTCTCTCC-----AGGAAGTGGAGTGGGGGTGGAAGCTCC 1251
```

```
Db 1304 ATTGAGCCAATAACTGCTGCTCCAGTGGCAGCGGAACGGGAGTAGACAGCTCT 1363
QY 1252 GGAGGAAGCTCTGGAGGAGTGGATCTGTGCGGACAAAGCAGATGGCTCTTACCCTGG 1311
Db 1364 GGAGGCACTCGGAGGAGTGGATCTGTGCTGTGAGAGCAACGGCCCTCTACCCCGTG 1423
QY 1312 CGAGATGACAGAAATGCTTTTGGCAGTGCATCAATGGAATCACATACCAAGCAGCATTTCT 1371
Db 1424 GCAATACAGAAATGCTTTTGGCAGTGCATCAATGGAATCACATACCAAGCAGCATTTCT 1483
QY 1372 CAAGCAGGCTTGTGTTTGTATACCAAGCTGTAAATGCTGCAACTGGCCTGAACCTAA 1428
Db 1484 CAGCGCGGCTGTCTGTGACACCAAGCTGTGATGCTGCAACTGGGCAATAAACCTGA 1540

RESULT 6
AB051629
LOCUS
DEFINITION
  Bos taurus cbp b04 mRNA for chitin binding protein b04, complete cds.
ACCESSION
  AB051629
VERSION
  AB051629.1 GI:16754894
KEYWORDS
  Bos taurus liver cDNA to mRNA.
ORGANISM
  Bos taurus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
  Bovidae; Bovinae; Bos.
REFERENCE
  1
  Suzuki,M., Morimatsu,M., Yamashita,T., Iwanaga,T. and Syuto,B.
  A novel serum chitinase that is expressed in bovine liver
  FEBS Lett. 506 (2), 127-130 (2001)
  2 (bases 1 to 1529)
  Suzuki,M., Morimatsu,M. and Syuto,B.
  Direct Submission
  Submitted (23-NOV-2000) Masako Suzuki, Iwate University, Department
  of Veterinary Medicine, Faculty of Agriculture, Ueda 3-18-8,
  Morioka, Iwate 020-8550, Japan (E-mail:u9998001@iwate-u.ac.jp,
  Tel:81-19-621-6212)
  Location/Qualifiers
    1..1529
     /organism="Bos taurus"
     /db_xref="taxon:9913"
     /tissue_type="liver"
    1..1529
     /gene="cbp b04"
    13..1431
     /gene="cbp b04"
     /codon_start=1
     /product="chitin binding protein b04"
     /protein_id="BAB71805.1"
     /db_xref="GI:16754895"
     /translation="MAKLIFLFLAFLNAQLGSAQLVLCVFSNMAQYRPGJGSKPKPD
     NIDPLCLHLIYAFAGMSNSEITTIEMNDVALYSFNLDKKNSQLKILATGGWNFG
     TAPFTAMATPENRKTFSVTKFLHOYDGLDFDWEYPGRSPGSPDKHLFTVLVQ
     ETRFAEQEAKTKNPRLLVTAAGAAGLSNIOAGYEIPOLSOYLDFIHWMTYDFHGSW
     EGYTGENSEPLKYPDTGTSNTLVNEYAMNWKNGAPAEKILIGFPAYGHNFILDA
     SNNGIGATSPGAGPYTRAGFWAYEICAFDKDGATEAWDDSDNVPYATKGTSEW
     GYDYNYSFRKIAQWLKNNFGGMYMAIDLDDTGTFCNOGKPEPLINTLKDGLKLSA
     TCNASTQSPGNSPNSGSGKSSSSSEBGRGCAGKADGLYPVADNRNRAFNCVNGI
     TVKONCLTGLVFDTSCHCNWA"
  polyA_signal
    1508..1513
     /gene="cbp b04"
  BASE COUNT
    391 a 405 c 365 g 368 t
  ORIGIN

Query Match 62.9%; Score 959.2; DB 4; Length 1529;
Best Local Similarity 80.0%; Pred. No. 1.5e-287;
Matches 1142; Conservative 0; Mismatches 283; Indels 3; Gaps 1;

QY 1 ATGGCCAGACTACTTCTCTGCTCAGAGCTCTGGCTCTTCTGCTGAATGCTCAGCTGGGGTCT 60
```



```

13 ATGCAAACTTATTTTCTTACAGGTCTGGCTTCTGCTGAATGCTCAGCTAGTCT 72
61 GCTTACAACTGATATGCTATTTTACCAGCTGGCCAGTATCGCCAGGCTCTGGGAGC 120
73 GCTTACAACTGATATGCTATTTTACCAGCTGGCCAGTATCGCCAGGCTCTGGGAGC 132
121 TTCAAGCCTGATGACATTAACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
133 TTCAAGCCTGATGACATTAACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 192
181 ATGCAAGCAATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240
193 ATGAGTAAACAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 252
241 AATGACTTGAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 300
253 AATGACTTGAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 312
301 TTTGGAAGCTGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 360
313 TTTGGAAGCTGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 372
361 ACCTCAGTCAATCAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
373 TCTTCAAGTCAATCAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 432
421 TACCCAGGCTCAGTGGGAGCTCTCTCAGGCAAGCAAGCAAGCAAGCAAGCAAGCA 480
433 TACCCAGGCTCAGTGGGAGCTCTCTCAGGCAAGCAAGCAAGCAAGCAAGCAAGCA 492
481 GAAATGCTGAGCTTTTGAAGCAAGCTTATGAGCAAGCAAGCAAGCAAGCAAGCA 540
493 GAAATGCTGAGCTTTTGAAGCAAGCTTATGAGCAAGCAAGCAAGCAAGCAAGCA 552
541 ACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
553 ACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 612
601 TCTAAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
613 TCCAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 672
661 TACACTGGGAGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
673 TACACAGGAGAAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 732
721 CTCATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
733 CTCATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 792
781 ATGTTGGATTTCCAGAGTATGACACACTTCTGAGCAAGCAAGCAAGCAAGCAAG 840
793 ATCATTTGATTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 852
841 ATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
853 ATCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 912
901 TGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
913 TGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 972
961 TCCCAAGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
973 TCTCAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1032
1021 AGCTTCAGTGTAAAGCTCAGTGGCTTAAAGCAAGCAAGCAAGCAAGCAAGCAAG 1080
1033 AGCTTCAGTGTAAAGCTCAGTGGCTTAAAGCAAGCAAGCAAGCAAGCAAGCAAG 1092
1081 GCCATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
```

```

1093 GCATTGACCTGGATGATTTTACAGGCACTTTTTCACACAGGCGCAAGTTCCCTCC 1152
1141 TCTACTTTGAACAAAGCCCTTTCATATCCACTGAAGTTGACAGCTCTCTGAGCTGC 1200
1153 AATACCTTGAAGGATGCTTTCCTGCTGAAGAGTGAACCTTGTATGCTTCCAGCAG 1212
1201 TCCAGCCAGTACTTCTTCCAGGAAGTGGAGTGGGGTGGAGTCTCCGAGGAGGAG 1260
1213 AGTGAGCCCAACAGCAGTCTCT---GGAACGAAAGTGGAGTGGAAATGAAGAGCAG 1269
1261 TCTGGAGGAGTGGATCTGTCGCAAGCAAGAGAGATGGCTTACCTGTGGCAGATG 1320
1270 TCCAGGGGTAGAGGATGCTGCTGCAAGAGTGGCTTACCTGAGGCGGAGTAAC 1329
1321 AGAAATGCTTTTGGCAGTGCATCAATGAATGCAATACAGCAGCAGATTTCAAGCAG 1380
1330 AGAAATGCTTTCGGAAGTGGCTGAATGCAATACAGCAGCAGATTTCAAGCAG 1389
1381 CTTGTTTTCATACCAAGCTGTAATGCTGCACTGGCCCTGAACCTAA 1428
1390 CTTGTTTTCATACCAAGCTGTAATGCTGCACTGGCCCTGAACCTAA 1437
```

RESULT 7

AB025008

LOCUS

DEFINITION

AB025008

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

gene

CDS

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches 1013;

Conservative

0;

Mismatches 230;

Indels 9;

Gaps 1;

Score 856;

DB 9;

Length 1354;

Pred. No. 2.5e-255;

311 a

387 c

348 g

308 t

NNAFWHCVNGVTYQOOCAGLVFDTSDDCCNWA

/organism="Homo sapiens"

/db_xref="taxon:9606"

/gene="TSA1902-L"

/product="novel member of chitinase family"

/protein_id="BA86980.1"

/translation="MVSTPENRFTTTSIKELROYEFDGLDFDWEYSGRSPPODK

HFTLVQEMREAFQEKQINKPRLMTVAAGVSIQSGYEIPQLSQYLDYIHVM

TYDLHWSWGYTGENSEPLYKPTDGSNAYLVNVDVYMYKNDNGAPAEKLIVGFPTG

HNFILSNPTSGTSGAGPAGPYKESGWAYVEICTLKNKGATOCGMDAPQEVY

AYQGNWGYDNVKSFDIKAWLKHKNKFGAMWALDDDFDTGTCNQGKPLISTLK

KALQASCTAPQAPFIEPTAIPSGSGSGSSSGSGSGSGSGSGSGSGSGSGSGSGSG

NNAFWHCVNGVTYQOOCAGLVFDTSDDCCNWA"

311 a

387 c

348 g

308 t

186 GAACAATGAGATCACCAACATAGAAATGGAATGATGTTACTCTCTATTAAGCTTTCAATGA 245
18 GCACGAACAGGTGCGCGACTCTGGAGCCAGCGTGTGCTTCCAGTCTGGTGAATC 77
246 CTTGAAAACAGGACAGCAACAACTGAAACCCCTCTGGCAATGAGGCTGGAATTTGG 305
78 CTCCATAGCTGGAAGCCAGCGCTGAAACCTCTCTGGCCATTTGAGGCTGGAATTTGAG 137
306 AACTGCTCTTCTACTACATGCTTCCACTTCTCAGAACCCGACAGCTTTCATACCTC 365
138 GACTGCCCTTCTACTGCCATGTTTCTACTCTTGAGAACCCGACAGCTTTCATACCTC 197
366 AGTCATCAATTTCTGGCTAGTATGGGTTTGTAGTGGACTGGACCTGGGATACCC 425
198 AGTCATCAATTTCTGGCCAGTATGAGTTTGACGGGCTGGACTTGGCTGGAGTACC 257
426 AGGCTCAGCTGGGAGCCCTCTCAGGACAAACATCTCTTCACTGCTCTGGTGAAGAAAT 485
258 TGGCTCTCGTGGGAGCCCTCTCAGGACAAACATCTCTTCACTGCTCTGGTGGCAGAAAT 317
486 GCGTGAAGCTTTTGGAGGAGGCTATTGAGCAACAGGCGCCAGACTGATGTTTACGTC 545
318 GCGTGAAGCTTTTGGAGGAGGCGCCAGCAGATCAACAGCCAGGCTGATGTTTACGTC 377
546 TGCTGAGCTGGTGGGATTTCCAAATCCAGGCTGGCTATGAGATCCCTGAACTTTCTAA 605
378 TGCAGTAGCTGCTGCATCTCCAATATCCAGTCTGGCTATGAGATCCCCCAACTGTACA 437
606 GTACTCGATTTTCATCATGTCATGACATGATGATGCTGCTGCTGGGAGGCTACAC 665
438 GTACTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 497
666 TGGGAGAAATAGTCTCTTTTCAAAATACCTTACTGAGACTGATGATGATGATGATGATGAT 725
498 TGGAGAGACAGCCCTCTCAAAATACCCAGTATGACCCGACGACGACGCTTACCTCAA 557
726 TGTGGATATGTCATGAACTATTGGAAGAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 785
558 TGTGGATATGTCATGAACTATTGGAAGAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 617
786 TGGATTTCCAGAGTATGAGACACATCTTCACTCTGAGAAACCCCTCTGATATGAAATGG 845
618 TGGATTTCCAGTATGAGACACATCTTCACTCTGAGAAACCCCTCTGAGAAATGG 677
846 TGCCCTTACCTCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 905
678 TGCCCTTACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 737
906 CTACTATGAGATTTGACCTTTCTGAGAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 965
738 TTTACTAGAGATCTGTACTTCTGAGAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 797
966 AGAAGTCCCTTACCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1025
798 GGAAGTCCCTTATGCTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 857
1026 CAGTGTAAAGCTCAGTGGCTTTAAGCAGAAATTTTGGAGGAGGAGGAGGAGGAGGAGGAG 1085
858 CGATATTAAGCTCAATGGCTTATGAGCAGAAATTTTGGAGGAGGAGGAGGAGGAGGAGGAG 917
1086 TGACCTTGAGACTTCACTGCTCTTTCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1145
918 TGATCTGGATGACTTCACTGCTCTTTCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 977
1146 TTTGAGCAAAACCCCTTGGCATAATGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1205
978 CCTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1037
1206 GCCAGTACTACTCTCTCC-----AGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGG 1256
1038 GCCAATACTCTCTCTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1097

QY 1257 AAGCTCTGGAGCAGTGGATTTCTGTGCGGACAAACAGATGGCCCTCTTACCCTGTGGCAGA 1316
Db 1098 CAGCTCGGAGGAGCAGTGGATTTCTGTGCGGAGCAACGCGCTCTTACCCTGTGGCAGA 1157
QY 1317 TGACAGAAATGCTTTTGGCAGTGCATCAATGGAATACATACACGAGCAGTGTGCAAGC 1376
Db 1158 TAACAGAAATGCTTTTGGCAGTGCATCAATGGAATGGAATGCACTACCGAGCAGAACTGCCAGGC 1217
QY 1377 AGGCTCTCTTTTGTATACCACTGTAATTTGTGCAACTGGCAGTGAACCTAA 1428
Db 1218 CGGCTGTCTTCGACACCACTGATGTTGCTGCACTGGGATTAACCTGA 1269
RESULT 8
AX405989
LOCUS AX405989 1369 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 404 from Patent WO0222660.
ACCESSION AX405989
VERSION AX405989.1 GI:21439417
KEYWORDS SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1
AUTHORS Tang, Y. T., Liu, C., Zhou, P., Asundi, V., Zhang, J., Zhao, Q. A., Ren, F.,
Xue, A. J., Yang, Y., Wehrman, T. and Drmanac, R. T.
TITLE Novel nucleic acids and polypeptides
JOURNAL Patent: WO 0222660-A 404 21-MAR-2002;
HYSEQ, INC. (US)
FEATURES
Location/Qualifiers
1..1369
/organism="Homo sapiens"
/db_xref="taxon:9606"
154..1260
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD35096.1"
/translation="MVSTPENROTFTISVIFKROYEDGLDFDWEYPSRSGSPQDK
HLFTVLQEMREAEQEKQKINPRLAWTAVAAGISINQSYEIPQLSQYLDYIHVM
TYDLHGSWEYTGNSPLIKYPTDTSNAYLVNDYVMYKNDGAPAKLIVGPEYTG
HNFILSNPNTGICATGSGAGPAGPYAKESGWAYEICTELKNGATGWDPAQEVY
AYQNVWVYDNIKSFIDKQWLNKFNFGAMVAILDDFTGFCNGKPLSLTK
KALQIASCTAPAPTEPIATAPSGNSGSSGSSGSSGSGGSCFCAVRANGLYPVAN
NRNFAFHCVNGVTYQNCQAGLVFDTSCDCNWA"
BASE COUNT 327 a 393 c 341 g 308 t
ORIGIN
Query Match 56.0%; Score 854; DB 6; Length 1369;
Best Local Similarity 82.6%; Pred. No. 1.1e-254;
Matches 993; Conservative 0; Mismatches 200; Indels 9; Gaps 1;
QY 236 CTTTCAATGACTTGAAAACAGGACAGCAACCTGAAAACCTCTCTGCAATTTGAGGCT 295
Db 65 CTTTACTGAAATTTGCAGCTCGGACAGCCAGCTGAAAACCTCTCTGCGCATTTGGAGGCT 124
QY 296 GGAATTTGGAACTGCTCTTTTCACTACCTACCTGTTTCCACTTCTCAGAACCCGACAGCT 355
Db 125 GGAATTTGGAACTGCTCTTTTCACTACCTACCTGTTTCTACTCTCTGAGAACCCGACAGCT 184
QY 356 TCATTAACCTCAGTCATCAAAATTTCTGCTCAGTATGCTGTTTGTAGTGGACTGGACCTGGACT 415
Db 185 TCATCACTCAGTCATCAAAATTTCTGCTCAGTATGCTGTTTGTAGTGGACTGGACTGGACT 244
QY 416 GGGAAATACCCAGGCTCAGTGGAGCCCTCTCAGGACAAAGATCTCTTCACTGTCCTGG 475
Db 245 GGGAGTACCTTGGCTCTCTGAGGAGCCCTCTCAGGACAAAGATCTCTTCACTGTCCTGG 304
QY 476 TGAGGAAATCGCTGAACCTTTTGGAGCAGGAGGCTTATTTGAGAGCAACAGGCCCCACACTGA 535
Db 305 TGCAGGAAATCGCTGAACCTTTTGGAGCAGGAGGCTTATTTGAGAGCAACAGGCCCCACACTGA 364


```
Db      1153 TCTACTTTAAAGGAGATCTCAATATACACAGTGCAGTTGCAAGGCGCTTA 1205
RESULT 10
D87757 1153 TCTACTTTAAAGGAGATCTCAATATACACAGTGCAGTTGCAAGGCGCTTA 1205
LOCUS   D87757 1153 TCTACTTTAAAGGAGATCTCAATATACACAGTGCAGTTGCAAGGCGCTTA 1205
DEFINITION Mus musculus mRNA for ECF-L precursor, complete cds.
ACCESSION D87757
VERSION   2
KEYWORDS  ECF-L precursor.
SOURCE    Mus musculus
ORGANISM  Mus musculus
REFERENCE 1 (bases 1 to 1506)
AUTHORS   Ohashi,M., Arita,H. and Hayai,N.
TITLE      Identification of a novel eosinophil chemotactic cytokine (ECF-L)
          as a chitinase family protein
JOURNAL   J. Biol. Chem. 275 (2), 1279-1286 (2000)
REFERENCE 2 (bases 1 to 1506)
AUTHORS   Ohashi,M.
TITLE      Direct Submission
JOURNAL   Submitted (06-SEP-1996) Makoto Ohashi, The University of
          Tokushima, Faculty of Integrated Arts and Sciences; 1-1
          Minami-Johsanjima, Tokushima, Tokushima 770, Japan
          (E-mail:ohashi@ias.tokushima-u.ac.jp, Tel:0886-56-7261,
          Fax:0886-56-7298)
COMMENT   On Oct 7, 1999 this sequence version replaced gi:1545818.
          Sequence updated (05-Oct-1999).
FEATURES             Location/Qualifiers
     source            1..1506
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
     gene              1..1506
                     /issue_type="bone marrow"
     CDS                1..1197
                     /gene="mECF-L"
                     /gene="mECF-L"
                     /codon_start=1
                     /product="ECF-L precursor"
                     /protein_id="BAAL3458.2"
                     /db_xref="gi:6015436"
                     /translation="MAKILVLTGLAILNVLGSSYQLMYYTSWAKDRPIEGSEKPG
                     NIDPLCTHLIVAPAGMONNETTYTHEODLYEALNGKDKNTKELKTLATLGGKFG
                     PASFAMYSTVONQIFQSVTRFLRQNFEDLNLDWOYPGSRGPPDKHLFSLVK
                     EMRAFESEVEKDIPLLSTGAGIDIVKSGYKIPELSOSLDYIQWYDLHDPK
                     DGYTGNSPLKSPYDIKGSADLNVDISIYWKDHGAASEKLIVGPAYGHTFILSDP
                     SKTGIGAPTISTGPGKYTDESGLAYEYCTFLNEGATEVMDAPQEPYAYQGNHW
                     GYDNRVSFKLKAQWLKDNLGVVWPLDMODFSGSFCHQRHPLTSLKGLDNLHSA
                     SCKGPY"
BASE COUNT 423 a 338 c 333 g 412 t
ORIGIN
Query Match 46.8%; Score 713; DB 10; Length 1506;
Best Local Similarity 74.9%; Pred. No. 1.2e-210;
Matches 893; Conservative 0; Mismatches 300; Indels 0; Gaps 0;
QY 1 ATGCGCAAGCTACTTCTCGTCACAGGCTGGCTCTTCTGCTGAATGCTCAGCTGGGGTCT 60
Db 1 ATGCGCAAGCTACTTCTGTCACAGGCTGGCAATTTCTTGAAGCTACAGCTGGGATCT 60
QY 61 GCCTACATCATGATGCTATTTCACCAACTGGGCCCCAGTATCGGCCAGGCTCGGGAGC 120
Db 61 TCCTACCAAGCTGATGCTGCTACTATATACAGTTGGGTAAAGGACAGGCCAATAGAAGGAGT 120
QY 121 TTCAAGCTGATGACATTAACCCCTGCCTGCTACTCAGCTCATGATCTATGCCCTTTCCTGG 180
Db 121 TTCAACCTGGTAATATTGACCCCTGCCTGCTACTCAGCTCATGATCTATGCCCTTTCCTGGA 180
QY 181 ATGCAGAAATAGATGATCACCACCAATAGGAATGATGTTACTCTCTATAAAGCTTTC 240
Db 181 ATGCAGAAATAGATGATCATTACACATGAGCAAGACTTGCCTGACTATGAAGCATTTG 240
```

```
Db      1153 TCTACTTTAAAGGAGATCTCAATATACACAGTGCAGTTGCAAGGCGCTTA 1205
RESULT 10
D87757 1153 TCTACTTTAAAGGAGATCTCAATATACACAGTGCAGTTGCAAGGCGCTTA 1205
LOCUS   D87757 1153 TCTACTTTAAAGGAGATCTCAATATACACAGTGCAGTTGCAAGGCGCTTA 1205
DEFINITION Mus musculus mRNA for ECF-L precursor, complete cds.
ACCESSION D87757
VERSION   2
KEYWORDS  ECF-L precursor.
SOURCE    Mus musculus
ORGANISM  Mus musculus
REFERENCE 1 (bases 1 to 1506)
AUTHORS   Ohashi,M., Arita,H. and Hayai,N.
TITLE      Identification of a novel eosinophil chemotactic cytokine (ECF-L)
          as a chitinase family protein
JOURNAL   J. Biol. Chem. 275 (2), 1279-1286 (2000)
REFERENCE 2 (bases 1 to 1506)
AUTHORS   Ohashi,M.
TITLE      Direct Submission
JOURNAL   Submitted (06-SEP-1996) Makoto Ohashi, The University of
          Tokushima, Faculty of Integrated Arts and Sciences; 1-1
          Minami-Johsanjima, Tokushima, Tokushima 770, Japan
          (E-mail:ohashi@ias.tokushima-u.ac.jp, Tel:0886-56-7261,
          Fax:0886-56-7298)
COMMENT   On Oct 7, 1999 this sequence version replaced gi:1545818.
          Sequence updated (05-Oct-1999).
FEATURES             Location/Qualifiers
     source            1..1506
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
     gene              1..1506
                     /issue_type="bone marrow"
     CDS                1..1197
                     /gene="mECF-L"
                     /gene="mECF-L"
                     /codon_start=1
                     /product="ECF-L precursor"
                     /protein_id="BAAL3458.2"
                     /db_xref="gi:6015436"
                     /translation="MAKILVLTGLAILNVLGSSYQLMYYTSWAKDRPIEGSEKPG
                     NIDPLCTHLIVAPAGMONNETTYTHEODLYEALNGKDKNTKELKTLATLGGKFG
                     PASFAMYSTVONQIFQSVTRFLRQNFEDLNLDWOYPGSRGPPDKHLFSLVK
                     EMRAFESEVEKDIPLLSTGAGIDIVKSGYKIPELSOSLDYIQWYDLHDPK
                     DGYTGNSPLKSPYDIKGSADLNVDISIYWKDHGAASEKLIVGPAYGHTFILSDP
                     SKTGIGAPTISTGPGKYTDESGLAYEYCTFLNEGATEVMDAPQEPYAYQGNHW
                     GYDNRVSFKLKAQWLKDNLGVVWPLDMODFSGSFCHQRHPLTSLKGLDNLHSA
                     SCKGPY"
BASE COUNT 423 a 338 c 333 g 412 t
ORIGIN
Query Match 46.8%; Score 713; DB 10; Length 1506;
Best Local Similarity 74.9%; Pred. No. 1.2e-210;
Matches 893; Conservative 0; Mismatches 300; Indels 0; Gaps 0;
QY 1 ATGCGCAAGCTACTTCTCGTCACAGGCTGGCTCTTCTGCTGAATGCTCAGCTGGGGTCT 60
Db 1 ATGCGCAAGCTACTTCTGTCACAGGCTGGCAATTTCTTGAAGCTACAGCTGGGATCT 60
QY 61 GCCTACATCATGATGCTATTTCACCAACTGGGCCCCAGTATCGGCCAGGCTCGGGAGC 120
Db 61 TCCTACCAAGCTGATGCTGCTACTATATACAGTTGGGTAAAGGACAGGCCAATAGAAGGAGT 120
QY 121 TTCAAGCTGATGACATTAACCCCTGCCTGCTACTCAGCTCATGATCTATGCCCTTTCCTGG 180
Db 121 TTCAACCTGGTAATATTGACCCCTGCCTGCTACTCAGCTCATGATCTATGCCCTTTCCTGGA 180
QY 181 ATGCAGAAATAGATGATCACCACCAATAGGAATGATGTTACTCTCTATAAAGCTTTC 240
Db 181 ATGCAGAAATAGATGATCATTACACATGAGCAAGACTTGCCTGACTATGAAGCATTTG 240
```


Db 241 AATGGTCTGAAGACAGGAACACACACAGTACAGTAAAAACTCTCTCGGCAATGGAGGATGGAAG 300
Qy TTTTGAAGTCTCTCTTTTACACACACATGATGTTTCCACTCTCTCAGAAACCGCAGACCTTCATT 360
Db TTTTGAAGTCTCTCTTTTACACACACATGATGTTTCCACTCTCTCAGAAACCGCAGACCTTCATT 360
Qy ACCCTGAGTCAATCAATTTCTCGTCACTATGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 420
Db AAGTCAGTATACAGATCTCTCGTCAATATCAATTTGATGGCTCAACCTGGAGTGGCAG 420
Qy TACCCAGGCTCAGCTGGGAGGCTCTCTCAGACAGACATCTCTTCACTGCTCTCTGGTGAAG 480
Db TACCTGGGCTCGAGGAGGCTCTCTTAAAGACAAACATCTCTTCACTGCTCTCTGGTGAAG 480
Qy GAAATGCTGAAGCTTTTGAAGCAGGAGCTATTGAGAGCAACAGCCAGACCTGATGTTT 540
Db GAAATGCTGAAGCTTTTGAAGAGAGATCTCACTTTGAACACATCTCAAGGCTGCTACTC 540
Qy ACTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
Db ACTTCCAGCAGGAGTGGATTCATTGACGTAATCAAGTCTGGGTACAAGATCCCTGAAGTGG 600
Qy TCTAAGTACCTGGATTCATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
Db TCTCAGTCTCTCGACTATATTCAGGCTCATGATGATGATGATGATGATGATGATGATGATGATG 660
Qy TACAGTGGGAGATAGTCTCTTTTACAAATACCTCTCTGAGAGCTGAGTGGTGAAGTGCCTAC 720
Db TACAGTGGGAGATAGTCTCTTTTACAAATACCTCTCTGAGAGCTGAGTGGTGAAGTGCCTAC 720
Qy CTCAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
Db CTCAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
Qy ATTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
Db ATTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
Qy ATGTTGCTGCTTACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Db ATAGTGTACCTACTGTTAGTGTGCTGGACCAACAGGAAGTACAAATGACAAAGGACTC 900
Qy TGGGCTACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
Db CTGGCTACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
Qy TCCCAAGAGTCCCTATGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
Db ACCCAGGAAGTACCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
Qy AGCTTCACTGTTAAGGCTCAGTGGCTTAAGCAGAACAAATTTTGGAGGTGCCATGATCTGG 1080
Db AGCTTCACTGTTAAGGCTCAGTGGCTTAAGCAGAACAAATTTTGGAGGTGCCATGATCTGG 1080
Qy GCCATGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
Db CCCCTGGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140
Qy TCTACTTTGACAAAGCCCTTGGCATATCCACTGAAGTTCGACAGCTTCCTGA 1193
Db ACTACTTTAAGAGAGATCTGAATGTACACAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGC 1193

RESULT 12

AY065557 LOCUS 1209 bp mRNA linear ROD 22-MAY-2002
DEFINITION Mus musculus putative secretory protein precursor (Ym2) mRNA,
complete cds.
ACCESSION AY065557
VERSION AY065557.1 GI:18086513
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1209)
Chang,N.-C.A.
Tissue-specific Expression of Ym2 Protein
Unpublished
2 (bases 1 to 1209)
Chang,N.-C.A.
Direct Submission
Submitted (05-DEC-2001) Institute of Microbiology & Immunology,
National Yang-Ming University, #155 Sec. 2 Li-Nong St., Beitou,
Taipei 112, Taiwan
Location/Qualifiers
1. 1209
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
/clone="pGEMT-ez-ym2.s"
/tissue_type="stomach"
1. 1209
/gene="Ym2"
1. 1209
/gene="Ym2"
/codon_start=1
/product="putative secretory protein precursor"
/protein_id="AAL57751.1"
/db_xref="GI:18086513"
/translation="MAKILIVTGLAILNVOLGSSYOLMCYVTSWAKDRPTGSGFKPG
NIDPLCLTHLIYAFAGMKNNEIYLSQDLRDYDALNGKDRNTEKTLKLLAIGKMKFG
PAFPSSMYVTPNQRTFKSVIRFLNQDFNLDMQVPGPRSPGPPKDLFLSVLQ
EMKAFEEESTLNHPRLLSTAGFIDVSKGYKIPELSQSLDYIQVMTYFLHDPK
NGYTGENSEPLXSPYDIGKSADLVDSIITWKDHGAASEKLIYGFAYGHTFILSDP
SKNGIDPTVSAGPPGYTNEGGLLAYFEICFELNEGATEIFDQTEVPYAVLGNWV
GYDNVRSEFKLKAQWLKNNLGGAVVWPLDMDDFSGSCFHQGRFPLTTLLKRLNVHSA
CKSKASYRGEI"

gene

CDS

BASE COUNT 326 a 274 c 282 g 327 t
ORIGIN

Query Match

Best Local Similarity 46.3%; Score 706.6; DB 10; Length 1209;
Matches 889; Conservative 0; Mismatches 304; Indels 0; Gaps 0;

Qy 1 ATGGCCAAAGCTACTTCTCGTCACAGGCTGGCTCTTCTGCTGAATGCTCAGCTGGGGTCT 60
Db 1 ATGGCCAAAGCTACTTCTTGTTCACAGGCTCTGGCAATCTTCTGAATGTACAGCTGGGATCT 60
Qy 61 GCCTACATCTGATATGCTATTTACCAACTGGGGCCAGATCGGCCAGGCTGGGGAGC 120
Db 61 TCCTACCAGCTGATGCTGCTACTATACAGCTGGGGCTAAGGACAGGCCAACAGAGGAGT 120
Qy 121 TTCAGCCCTGATGACATTAACCCCTGCTGTACTCACCTGATCTATGCCCTTCTGCTGG 180
Db 121 TTCAGCCCTGATGATTAATTTGACCCCTGCTGTACTCACCTGATCTATGCCCTTCTGCTGG 180
Qy 181 ATGCAGAAACATAGATCACCACCATAGAAATGATGTTACTCTCTATAAAGCTTTC 240
Db 181 ATGAAGAATAATGAGATCATTACTTAAAGTGAGCAAGACTTCCGTGACTATGAAGCAATTA 240
Qy 241 AATGACTTGAACAAACAGCAACAGCAAACTGAAACCCCTCTGGCAATTTGGAGGCTGGAC 300
Db 241 AATGCTTGAAGACAGCAACAGCAAACTGAAACCCCTCTGGCAATTTGGAGGATGGAG 300
Qy 301 TTTGGAACCTGCTCTCTTTCACCTACCATTGTTTCCACTTCTCAGAACCCGACAGACCTTCATT 360
Db 301 TTTGGAACCTGCCCTGCTCAGTTCCATGTTCTCTACTCTCAGAACCCGACAGACCTTCATT 360
Qy 361 ACCTCAGTCATCAAAATTTCTGCGTCAGTATGGTGTGATGGACTGGACCTGGACCTGGGAA 420
Db 361 AAGTCAGTTATCAGATCTCTTCTCGTCAATATTAATTTGATGGCTCAACCTGGACCTGGCAG 420
Qy 421 TACCAGGCTCAGCTGGGAGGCTCTCTCAGGACAGCAATCTCTTCACTGTCTCTGTGTGAAG 480
Db 421 TACCCTGGGCTCGAGGAGGCTCTCTTAAGGACAAACATCTCTTCACTGTCTCTGTGTGAG 480

QY 481 GAAATCGTGAAGCTTTTGAGCAGGAGGCTATTGAGAGCAACAGGCCAGACTGATGGTT 540
|||||
Db 481 GAAATCGTGAAGCTTTTGAGGAGGAATCCACATTTGAACACATTTCCAGGCTGCTACTC 540
QY 541 ACTGCTGCTGAGCTGGTGGATTTCCACATCCAGGCTGGCTGATGAGATCCCTGAACTT 600
|||||
Db 541 ACTTCCACAGGAGCTGGATTCATTTGACGTAATCAAGTCTGGGTACAGATCCCTGAACTG 600
QY 601 TCTAAGTACTGATTTCTATCCATGTCATGACATATGACATCCATCCATGCTCTGGAGGCG 660
|||||
Db 601 TCTAAGTACTGATTTCTATCCATGTCATGACATATGACATCCATCCATGCTCTGGAGGCG 660
QY 661 TACACTGGGAGAAATAGTCTCTTTTACAAATACCCTACTGAGACTGGTAGCAATGCTAC 720
|||||
Db 661 TACACTGGGAGAAATAGTCTCTCTTATAATCTCCATATGACATTTGGAAGAGTCTGAT 720
QY 721 CTCATGTTGGATTTATGATGATGATTTGGAAGAAATGAGGCCGCCAGCTGAGAGCTC 780
|||||
Db 721 CTCATGTTGGATTTATGATGATGATTTGGAAGAAATGAGGCCGCCAGCTTCTGAGAGCTC 780
QY 781 ATTGTTGGATTTCCAGAGTATGACACACACCTTCTGAGAAACCCCTCTGATAATGGA 840
|||||
Db 781 ATTGTTGGATTTCCAGAGTATGATGATGATTTCTGAGTACCCCTTCTGATAATGGA 840
QY 841 ATTGTTGGCTCCCTACCTCTGATGATGATGATGATGATGATGATGATGATGATGATG 900
|||||
Db 841 ATAGTGTACCTACTGTTAGTGTGACCAACAGGAGGATGACAAATGAACAAGGACTC 900
QY 901 TGGGCTCTACTATGAGATTTGACCTTTCTGAGAAAGTGGAGCCACTGAGGTCTGGATGCC 960
|||||
Db 901 CTGGCTTACTTTGAGATTTGATGATTTCTGAGTGAAGAGGACCTGAGATCTTTGATGCC 960
QY 961 TCCAAAGAGTGGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
|||||
Db 961 ACCAGGAGTACCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
QY 1021 AGCTTCAAGTGAAGGCTGAGTGGCTCAAGGAGCAACAATTTAGGAGGTGGCTGGTGG 1080
|||||
Db 1021 AGCTTCAAGTGAAGGCTGAGTGGCTCAAGGAGCAACAATTTAGGAGGTGGCTGGTGG 1080
QY 1081 GCCATTGACCTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140
|||||
Db 1081 CCCTTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140
QY 1141 TCTACTTTGAACAAAGCCCTTGATGATGATGATGATGATGATGATGATGATGATGATG 1193
|||||
Db 1141 ACTACTTTAAAGAGAGATCTGAATGATGATGATGATGATGATGATGATGATGATGATG 1193

RESULT 13

AK098814
LOCUS
DEFINITION
Homo sapiens cDNA FLJ25948 fis, clone STM04207, highly similar to
Homo sapiens acidic mammalian chitinase precursor.
AK098814
ACCESSION
VERSION
KEYWORDS
SOURCE
oligo capping; fis (full insert sequence).
Homo sapiens stomach mucosa cDNA to mRNA, clone_lib:STM
clone:STM04207.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1
Ninomiya, K., Wagatsuma, M., Kanda, K., Kondo, H., Yokoi, T.,
Kodaira, H., Furuya, T., Takahashi, M., Kikkawa, E., Omura, Y., Abe, K.,
Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M.,
Suzuki, Y., Hata, H., Nakagawa, K., Mizuno, S., Morinaga, M.,
Kawanura, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H.,
Nishikawa, T., Sugiyama, A., Kawakami, B., Nagai, K., Isogai, T. and
Sugano, S.

NEDO human cDNA sequencing project

Unpublished

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

2 (bases 1 to 4250)
Sugano, S. and Suzuki, Y.
Direct Submission
Submitted (08-JUL-2002) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure, Human Genome
Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
(E-mail: cdna@ims.u-tokyo.ac.jp, Tel: 81-3-5449-5286,
Fax: 81-3-5449-5416)

COMMENT

NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction and 5'-end one pass sequencing: Institute of Medical
Science, University of Tokyo, Laboratory of Genome Structure, Human
Genome Center; 3'-end one pass sequencing: RAB; clone selection for
full insert sequencing: RAB and Helix Research Institute.

FEATURES

source

1..4250
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="STM04207"
/tissue_type="stomach mucosa"
/clone_lib="STM"
/note="cloning vector: pME18SFL3"

BASE COUNT

1262 a 868 c 966 g 1154 t

ORIGIN

Query Match 43.9%; Score 669; DB 9; Length 4250;
Best Local Similarity 81.2%; Pred. No. 7.8e-197;
Matches 793; Conservative 0; Mismatches 175; Indels 9; Gaps 1;

QY 461 TCTTCTACTGCTCCGTGGAAGGAATCGGTGAAGCTTTTGAGCAGGAGGCTTATTGAGAGCA 520
|||||
Db 3184 TATGCTCTTATTATTTCTGTAGGAATCGTGAAGCTTTTGAGCAGGAGGCGCAAGCATCA 3243
QY 521 ACAGGCCAGACTGATGTTTACTGCTGCTGAGGATTTTCCACATCCAGGCTG 580
|||||
Db 3244 ACAGGCCAGGCTGATGTTTACTGCTGCTGAGGATTTTCCACATCCAGGCTG 3303
QY 581 GCTATGAGATCCCTGAACTTTCTTAAGTACCTGGATTTTCATCCATGATGATGATGATG 640
|||||
Db 3304 GCTATGAGATCCCTGAACTTTCTTAAGTACCTGGATTTTCATCCATGATGATGATG 3363
QY 641 TCCATGGCTCCCTGGAGGCTACACTGGGAGATAGTCTCTTTTCAATACCTACTG 700
|||||
Db 3364 TCCATGGCTCCCTGGAGGCTACACTGGGAGATAGTCTCTTTTCAATACCTACTG 3423
QY 701 AGACTGGTAGCAATGCTTACCTCAATGTGGATTTATGATGAATTTTGAAGAACTG 760
|||||
Db 3424 ACACCGCAGCAAGCTTACCTCAATGTGGATTTATGATGAATTTTGAAGAACTG 3483
QY 761 GAGCCCCAGCTGAGAGCTCATTTGTTGGATTTCCAGAGATGATGACACACCTTCCATCTGA 820
|||||
Db 3484 GAGCACCAGCTGAGAGCTCATGTTGGATTTCCCTACCTATGAGCACAACCTTCCATCTGA 3543
QY 821 GAAACCCCTCTGATTAATGGAATTTGGTGGCCCTACCTCTGTTGGTGGCCCTGCTGGCGCT 880
|||||
Db 3544 GCAACCCCTCCAAACACTGGAATTTGGTGGCCCTACCTCTGTTGGTGGCCCTGCTGGCGCT 3603
QY 881 ATACCAAGACAGGCTGGGTTCTGGCCCTACTATGAGATTTTGCACCTTTCTGAGAAGTGGAG 940
|||||
Db 3604 ATGCCAAGAGCTGAGGATCTGGGATCTGGGCTTACTACGAGATCTGTACCTTCTGAGAAATGGAG 3663
QY 941 CCAGTGGTCTGGGATGCTCCCAAGAGTGGCCCTATGCTATGAGGCGCAACAGTGGC 1000
|||||
Db 3664 CCAGTGGGATGGGATGCTCCCTCAGGAAGTGGCTTATGCTATCAGGCGCAATGTGTGGG 3723
QY 1001 TTGGCTTATGACATATCAAGAGCTTTCAGTCTTAAGGCTCAGTGGCTTAAGCAGAACT 1060
|||||
Db 3724 TTGGCTTATGACATATCAAGAGCTTTCAGTCTTAAGGCTCAGTGGCTTAAGCAGAACT 3783
QY 1061 TTGGAGTGGCTATGCTGGGCTTACCTTGTGATGCTTCTGCTGCTGCTGCTGCTGCTG 1120
|||||
Db 3784 TTGGAGGCGCCATGGTCTGGGCTTGTGATGCTGATGCTGCTGCTGCTGCTGCTGCTG 3843

[illegible]


```
VERSION U56900.1 GI:1336165
KEYWORDS
SOURCE house mouse strain-CBA.
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 994)
AUTHORS Shmelkov,S.V., Zinovjeva,M.V. and Belyavsky,A.V.
TITLE Mouse chitinase-related protein mRNA (MCRP), partial cds
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 994)
AUTHORS Shmelkov,S.V.
JOURNAL Direct Submission
TITLE Submitted (27-APR-1996) Sergey V. Shmelkov, Molecular Basis of
Differentiation and Development, Engelhardt Institute of Molecular
Biology, Vavilov 32, Moscow 117984, Russia
FEATURES
    source
        1..994
            /organism="Mus musculus"
            /strain="CBA"
            /db_xref="taxon:10090"
            /tissue_type="bone marrow"
            <1..784
                /codon_start=2
                /product="chitinase-related protein MCRP"
                /protein_id="AA01230.1"
                /db_xref="GI:1336166"
                /translation="WVPGSRGPPKDKHLSVLKEMKKAFFEESEVKDIPRLILTS
                TGALIDIVKSGYKIPELSLDYIOVMYDLHDPKDGTYGENSPLYKSPYDYGKSNAD
                LNVDSIISWKHGKASEKLIVGPAITFTILSDPSKTIAGTISTGPPGKIDTES
                GLIAYEVCTFLNEGAEVMDAFOEYAYQGNWYGVNDVRSFKLKAQWLKDNLLGG
                AVYPLMDDFSGSFQRHFLPTLTILHSDKSGPY"
                278 a 218 c 230 g 267 t 1 others
BASE COUNT
ORIGIN
Query Match 30.2%; Score 460; DB 10; Length 994;
Best Local Similarity 74.4%; Pred. No. 1.1e-131;
Matches 580; Conservative 0; Mismatches 200; Indels 0; Gaps 0;
QY 414 CTGGGAATACCCAGGCTCAGTGGAGCCCTCTCAGGCAAGCATCTCTTCACTGTCCT 473
DB 1 CTGGCAGTACCCCTGGGTCTCGAGAGCCCTCTCAGGCAAGCATCTCTTCACTGTCCT 60
QY 474 GGTGAAGAAATCGTGAAGCTTTTGAGCAGGAGGCTATTGAGAGCAACAGGCCCGAGCT 533
DB 61 GGTGAAGAAATCGTGAAGCTTTTGAGGAGAAATCTGTGAGAAAGACATTCGAAGGCT 120
QY 534 GATGGTTACTGCTGCTGTAGCTGGTGGGATTTTCAACATCCAGCTGGCTGATGAGATCCC 593
DB 121 GCTACTACATTCACAGGAGCAGGAATCATTTGACGTAATCAAGTCTGGCTACAGATCCC 180
QY 594 TGAACCTTCTAAGTACCTGATTTCCATGTCATGACATATGACCTCCATGGCTCCTG 653
DB 181 TGAACCTGCTCAGTCTCTGACTATATTTCAGGTCATGACATATGATCTCCATGATCCTAA 240
QY 654 GGAGGCTACACTGGGAGAAATAGTCCCTTTTAAATACCCCTACTGAGACTGGTAGCAA 713
DB 241 GGATGGCTACACTGGGAAATAGTCCCTCTTAAATCTCCATATGACATTTGAAAGAG 300
QY 714 TGCCCTACCTCAATGTGGATTATGTCATGAACATATTGGAAGAAATGGAGCCCGAGCTGA 773
DB 301 TGCTGATCTCAATGTGGATTCAATTTCTTACTGGAAGGACCATTGGAGCAGCTTCTGA 360
QY 774 GAAGCTCATTTGTTGGATTCCAGAGTATGGACACACTTCCATCTGAGAAACCCCTCTGA 833
DB 361 GAAGCTCATTTGTTGGATTTCACCATATGGGCATACCTTTATCTGAGTGACCCCTCTAA 420
QY 834 TAATGGAATTTGGTCCCTACCTCTCTGGTATGGCCCTGCTGGCGCCCTATACACAGAGGC 893
DB 421 GACTGGAATTTGGTCCCTACATATTAGTACTGGCCCGCAGGAAAGTACACAGATGAATC 480
QY 894 TGGGTTCTGGGCTACTATGAGATTTGCACCTTTTCTTCTGAGAAAGTGGAGGCGCTAGTCTG 953
    . . .
```

Search completed: July 3, 2003, 07:21:40
Job time : 4013.96 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 2, 2003, 22:31:11 ; Search time 349.54 Seconds
(without alignments)
9825.200 Million cell updates/sec

Title: US-10-004-219B-3

Perfect score: 1525

Sequence: 1 atggcagactactctctgt.....taaaattgtagccaaaaca 1525

Scoring table: IDENTITY_NUC

Gap 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002.*

- 1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
- 2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
- 3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
- 4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
- 5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
- 6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
- 7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
- 8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
- 9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
- 10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
- 11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
- 12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
- 13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
- 14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
- 15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
- 16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
- 17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
- 18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
- 19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
- 20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
- 21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
- 22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
- 23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
- 24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1042.6	68.4	1678	22 AAH42025	Disease treatment
2	999.2	65.5	1368	22 AAH42013	Disease treatment
3	854	56.0	1369	24 ABN59993	Novel human coding
4	714.6	46.9	1469	22 AAH42023	Disease treatment
5	426.8	28.0	1636	19 AAV10435	Human chitinase cl
6	426.8	28.0	1636	20 AAZ21847	MO-218 clone of hu
7	426.8	28.0	1636	22 AAD03759	Human chitinase cd
8	426.8	28.0	1636	24 ABL57380	Human chitinase cd
9	426.8	28.0	1643	18 AAT50833	Human chitinase cd

10	426.8	28.0	1713	18 AAT50834	Human chitinase cd
11	426.4	28.0	1768	18 AAT89181	Human chitinase cl
12	425.2	27.9	1656	19 AAV10436	Human chitinase cl
13	425.2	27.9	1856	20 AAZ21848	MO-13B clone of hu
14	425.2	27.9	1856	22 AAD03760	Human chitinase cd
15	425.2	27.9	1656	24 ABL57381	Human chitinase cd
16	406.4	26.6	1637	18 AAT89180	Human chitinase cl
17	362	23.7	2504	16 AAQ90443	Murine oviduct spe
18	350.8	23.0	2366	16 AAQ90444	Hamster oviduct sp
19	345.2	22.6	638	24 ABK11713	DNA encoding novel
20	344.8	22.6	746	22 AKR86635	Human digestive sy
21	343.2	22.5	1432	19 AAV21689	DNA encoding a hum
22	343.2	22.5	1433	19 AAV13925	Human cartilage gp
23	339.4	22.3	1994	16 AAQ90442	Bovine oviduct spe
24	338.8	22.2	1496	18 AAT97127	Human cartilage gl
25	335.6	22.0	1594	19 AAV21688	DNA encoding a hum
26	330.8	21.7	1526	19 AAV13926	Human cartilage gp
27	330.8	21.7	1594	19 AAV21687	DNA encoding a hum
28	327.6	21.5	1391	24 AAS62615	CDNA sequence #402
29	314.8	20.6	1681	16 AAQ85245	YKL-40 gene. Homo
30	314.8	20.6	1681	21 AA294901	Human cancer marke
31	307.4	20.2	1152	18 AAT99452	Bovine whey protei
32	296.6	19.4	2346	23 ABL03597	Drosophila melanog
33	264.4	17.3	17368	23 ABL03596	Drosophila melanog
34	259.8	17.0	1474	22 AAH23078	Osteoarthritis tis
35	210	13.8	966	9 AAN81756	Gene encoding poly
36	176.2	11.6	13497	23 ABL03557	Drosophila melanog
37	152.6	10.0	503	21 AAC00233	Human secreted pro
38	151.6	9.9	288	24 ABN76932	Human glycoprotein
39	146	9.6	2089	23 ABL29841	Drosophila melanog
40	142.2	9.3	3899	23 ABL19747	Drosophila melanog
41	140	9.2	2452	18 AAT62557	Manduca sexta larv
42	139.8	9.2	239	21 AAA42054	Human secreted exp
43	137.8	9.0	369	22 AAH42024	Disease treatment
44	136	8.9	1433	23 ABL15099	Drosophila melanog
45	132.6	8.7	1302	23 ABL19745	Drosophila melanog

ALIGNMENTS

RESULT 1
AAH42025
ID AAH42025 standard; DNA; 1678 BP.
AC AAH42025;
XX
XX 24-AUG-2001 (first entry)
DT Disease treatment related oligonucleotide SEQ ID NO: 16.
DE Disease treatment: infection; chronic occlusive pulmonary disease;
KW bronchial asthma; ds.
XX Homo sapiens.
XX WO200136633-A1.
XX 25-MAY-2001.
XX 14-NOV-2000; 2000WO-JP08015.
XX 15-NOV-1999; 99JP-0324467.
XX (TAKE) TAKEDA CHEM IND LTD.
XX Nakanishi A, Morita S;
XX WPI; 2001-397791/42.
XX New proteins, peptides and DNA for treatment of bronchial asthma,
XX chronic occlusive lung disease and infectious disease -

PS Example 5; Page 107-108; 114pp; Japanese.

The present invention provides the sequence of a protein which can be used in the treatment and prevention of infectious diseases. Inhibitors of the protein can be used to treat bronchial asthma and chronic obstructive pulmonary disease. The present sequence is an oligonucleotide described in the exemplification of the invention.

Sequence 1678 BP: 397 A: 497 C: 418 G: 366 T: 0 other: XX

Query Match 68.4%; Score 1042.6; DB 22; Length 1678;
Best Local Similarity 83.4%; Pred. No. 5.4e-307;
Matches 1199; Conservative 0; Mismatches 229; Indels 9; Gaps 1;

Qy	1	ATGCCAAGCTACTTCTCGCTACAGGTCTTGGCTCTTCTGCTGAATGCTCAGCTGGGGTCT	60
Db	143	ATGACAAGCTTATTTCTCCTCACAGGTCTTGTCTCTTATCTACTGAATTTGCAGCTCGGCTCT	202
Qy	61	GCCTACAACTGTATGATCTATTTACCAACTGGGCCAGTATCGGCCAGGTCTGGGGAGC	120
Db	203	GCCTACAGCTGACATGCTTACTTCCAACTGGGCCAGTACCGGCCAGGCTGGGGCGC	262
Qy	121	TTCAAGCCTGATGACATTAACCCCTGCCTGTGTACTCACCTGATCTATGCTTTGCTGGG	180
Db	263	TTCATGCTGACAACATCGACCCCTGCTGTGTACCCACCTGATCTACGCTTTGCTGGG	322
Qy	181	ATCGAGAACAAATGAGATCACCACTACCAATAGAAATGGAATGATTTACTCTCTATAAGCTTTC	240
Db	323	AGGCAGAACACGAGATCACCACTCGAAATGGAATGATGTGACTCTCTACCAAGCTTTC	382
Qy	241	AATGACTTGAAAACAGGAACAGCAAACTGAAAACCCCTCCTGGCAATTTGAGAGCTGGAAC	300
Db	383	AATGGCCTGAAAATAAGAACAGCCAGCTGAAAACCTCTCTGGCCATTTGAGGCTGGAAC	442
Qy	301	TTTGGAACTGCTCTTTCACCTACATGGTTTCCACTTCTCAGAACCCGACAGCTTCATT	360
Db	443	TTGGGACTGCCCTTTTCACTGGCATGGTTTCTACTCTGAGAACCCGACAGCTTTGATC	502
Qy	361	ACCTCAGTCATCAAAATTTCTGCGTTCAGTATGGGTTTGATGACTGTGACCTGGACTGGAA	420
Db	503	ACCTCAGTCATCAAAATTTCTTGGCCAGTATGAGTTTGACGGCTGGACTTTGACTGGAG	562
Qy	421	TACCGAGGCTCAGTGGGAGCCCTCTTCAGGACAAGCATCTCTTCACTGTCTGGTGAAG	480
Db	563	TACCTTGGCTCTCGTGGGAGCCCTCTCAGGACAAGCATCTCTTCACTGTCTGGTGCAG	622
Qy	481	GAATGGCTGAAGCTTTTGAGCAGGAGGCTATTGAGAGCAACAGGCCACACATGATGGTT	540
Db	623	GAATGGTGAAGCTTTTGAGCAGGAGGCCACAGCATCAACAAGCCAGCTGTATGGTC	682
Qy	541	ACTGCTGCTAGCTGGTGGGATTTCCAACTCCAGCTGGCTGATGAGATCCCTGAACTT	600
Db	683	ACTGCTGCAGTAGCTGTGGCATCTCCAATATCCAGTCTGGCTATGAGATCCCCAACTG	742
Qy	601	TCTAAGTACTGGATTTCACTCCATGTCATGACATATGACCTCCATGGCTCTGGGAGGC	660
Db	743	TCACAGTACCTGACATACATTCATGTCATGACCTTACGACCTCCATGGCTCTGGGAGGC	802
Qy	661	TACACTGGGAGAAATAGTCTCTTTCAAATACCTTACTCAGACTGGTAGCAATGCCTAC	720
Db	803	TACACTGGAGAGAAGCCCCCTCTACAATATCCCGACTCACACCGGCACACGCTTAC	862
Qy	721	CTCAATGTGGATTTATGTTCATGAATTTTGAAGAACAATGGAGCCCGACTGAGAAGCTC	780
Db	863	CTCAATGTGGATTTATGTTCATGAATCTTGAAGGACAATGGACCCAGCTGAGAAGCTC	922
Qy	781	ATTGTTGGATTTCCAGAGTATGACACACCTTCACTCTGAGAAACCCCTCTGATATGGA	840
Db	923	ATCGTTGGATTCCTTACCTATGACACAACCTTCACTCTGAGCAACCCCTCCACACCTGA	982
Qy	841	ATTGGTGGCCCTTACCTCTTGGTGAATGGCCCTGCTGGGCCCTTATACCAAGACAGCTGGTTC	900
Db	983	ATTGGTGGCCCACTCTGTGTGCTGTGCTGTGCTGGGCCCTTATGCCAAGGAGTCTGGATC	1042

Db 9 ATGCCAGCTCATCTTCTGTGTCACAGGCTGTGCAATCTCTCTGAAGCTACAGCTGGGATCT 68
QY 61 GCCTACATCTGATATGCTATTTTACCACACTGGGCCAGTAGTATCGCCAGGCTCTGGGAGC 120
Db 69 TCCCTACACAGCTGATGTGTCTATATACAGTGGGCTAAGGACAGGCCAATAGAAGGGAGT 128
QY 121 TTCAAGCCTGATGACATTAACCCCTGCTGTGTACTACCTGATGATCTATGCTCTTGTCTGG 180
Db 129 TTCAACACTGATATGACCCCTGCTGTGTACTACCTGATGATCTATGCTCTTGTCTGG 188
QY 181 ATGCAGAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240
Db 189 ATGCAGAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 248
QY 241 AATGACTTGAAGAAAGGACAGCAAACTGGAACCCCTGCGCAATTTGGAGGCTGGAAAC 300
Db 249-AATGGCTGAAGAAAGGACAGCAAACTGGAACCCCTGCGCAATTTGGAGGCTGGAAAC 308
QY 301 TTTGAACTGCTCTTTTACCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
Db 309 TTTGAACTGCTCTTTTACCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 368
QY 361 ACCTGATGATCAAAATTTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 420
Db 369 CAGTCAGTTATCAGATTCCTTCGTCATATTAATTAATTAATTAATTAATTAATTAATTA 428
QY 421 TACCAGGCTCAGTGGAGGCTCTCTCAGCAGCAATCTCTCAGTCTCTCTCTCTCTCTCTCT 480
Db 429 TACCCTGGGCTCGAGGAAGGCTCTTAAAGCAAACTCTCTCAGTCTCTCTCTCTCTCTCT 488
QY 481 GAAATGCTGAGGCTTTTGGAGGAGGCTATTTAGAGCAACAGGCCAGGCTGATGCTT 540
Db 489 GAATGCTGAAGCTTTTGGAGGAGGCTATTTAGAGCAACAGGCCAGGCTGATGCTT 548
QY 541 ACTGCTGCTGATGCTGGGATTTTCCAACTCCAGGCTGCTATGAGATCCCTGGAACCTT 600
Db 549 ACTTCCAGGAGGAGGAATCATTCAGCTAATCAAGTCTGGGTCAAGATCCCTGAAGCTG 608
QY 601 TCTAGTACCTGGATTCATCCATGCTCATGATGATGATGATGATGATGATGATGATGATGAT 660
Db 609 TCTAGTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 668
QY 661 TACACTGGGAGATAGTCTCTTTACAAATACCTCTGAGACTGGTAGCAATCCCTAC 720
Db 669 TACACTGGGAGATAGTCTCTTTACAAATACCTCTGAGACTGGTAGCAATCCCTAC 728
QY 721 CTCAATGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
Db 729 CTCAATGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 788
QY 781 ATGTTGGATTTCCAGAGTATGGACACACTTCATCTCTGAGAAACCCCTCTGATATGGA 840
Db 789 ATGTTGGATTTCCAGAGTATGGGATATGATGATGATGATGATGATGATGATGATGATGAT 848
QY 841 ATGTTGGCTTACCTCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Db 849 ATGTTGGCTTACCTCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 908
QY 901 TGGGCTTACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
Db 909 CTGGCTTACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 968
QY 961 TCCCAAGAGTCCCTATGCTTATGAGCCACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
Db 969 CCCCAGGAAGTACCTATGCTTATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1028
QY 1021 AGCTTCACTGTTAAGGCTCAGTGGCTTAAGCAGAACAAATTTGGAGGCTGCCATGATCGG 1080
Db 1029 AGCTTCACTGTTAAGGCTCAGTGGCTCAGGACACAAATTTAGGAGGCTGCCATGATCGG 1088
QY 1081 GCCATGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140

Db 1089 CCCCTGGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1148
QY 1141 TCTACTTTTGAACAAAGCCCTTGGCATATCCACTGAAGTTGTCACAGCTCCCTGA 1193
Db 1149 TCTACTTTTAAAGGAGATCTCAATATACAGTGCAGTTGCAAGGGCCCTTA 1201
RESULT 5
AAV10435
ID AAV10435 standard; cDNA; 1636 BP.
XX AAV10435;
AC AAV10435;
DF 15-JUN-1998 (first entry)
XX Human chitinase clone MO-218 cDNA.
DE Chitinase; human; fungal infection; immunogen; diagnosis; treatment;
KW Gaucher's disease; transgenic; detection; hybridisation; antifungal;
KW rheumatoid arthritis; overexpression; extracellular matrix; ss.
XX Homo sapiens.
XX Key Location/Qualifiers
FH CDS 2..1402
FT sig_peptide /*tag= a
FT /*tag= b
FT mat_peptide 65..1399
FT /*tag= c
FT /product= chitinase
FT /note= "from clone MO-218"
XX WO9747752-A1.
XX 18-DEC-1997.
XX 16-JUN-1997; 97WO-US10460.
XX 14-JUN-1996; 96US-0663618.
XX (ICOS-) ICOS CORP.
XX Gray PW;
XX WPI: 1998-052316/05.
XX P-PSDB; AAW40259.
XX Nucleic acids encoding human chitinase - useful as antifungal
agents, especially in combination with other antifungals
XX Claim 3; Page 38-40; 63pp; English.
XX This sequence encodes a novel human chitinase isolated from clone MO-218.
Chitinases are useful for treating or preventing fungal infection and
as immunogens for generating antibodies which are used to purify, detect
and quantify chitinases, e.g. for diagnosis of Gaucher's disease. The
nucleic acid sequence of the chitinase is also useful as a probe to
identify and isolate genomic DNA encoding chitinases or similar proteins,
or cells expressing them or to generate transgenic ('knockout') rodents.
It can also be used in hybridisation assays and to detect genetic
alterations in the chitinase gene related to disease. Agents that inhibit
this protein may be useful in treatment of Gaucher's disease and
rheumatoid arthritis, where overexpression of the protein can damage
the extracellular matrix. Chitinase also improves the activity of other
antifungal agents and may allow a reduction in the dose of such agents,
and thus of their side effects.
XX Sequence 1636 BP; 361 A; 491 C; 440 G; 344 T; 0 other;
Query Match 28.0%; Score 426.8; DB 19; Length 1636;
Best Local Similarity 62.1%; Pred. No. 3.4e-119;
Matches 693; Conservative 0; Mismatches 417; Indels 6; Gaps 1;

CC which are expected to be useful as products for treating fungal
CC infections or for the development of such products. Human
CC chitinase has a synergistic effect on the actions of other
CC fungicides. It can be administered to improve the antifungal
CC activity of a non-chitinase antifungal agent, especially
CC amphotericin B or itraconazole, in the treatment of a fungal
CC infection such as candidiasis, aspergillosis, coccidioidomycosis,
CC blastomycosis, paracoccidioidomycosis, histoplasmosis,
CC cryptococcosis, chromoblastomycosis, sporotrichosis, mucormycosis,
CC dermatophytoses and Pneumocystis infections (all claimed). In
CC particular, the fungal infection involves candida, Aspergillus
CC and/or Cryptococcus spp., whose growth is not effectively
CC inhibited by contact with human chitinase alone.
XX
XX
Query Match 28.0%; Score 426.8; DB 24; Length 1636;
Best Local Similarity 62.1%; Pred. No. 3.4e-119;
Matches 693; Conservative 0; Mismatches 417; Indels 6; Gaps 1;
QY 23 CAGGTCTGGCTCTTCTGCTGAATGCTCAGCTGGGCTGCGCTACAATCTGATATCTATT 82
DB 24 CAGGTCTGGCTCTTCTGCTGATGATCCATGGGCTCTGCTGCAAACTGGTCTGCTACT 83
QY 83 TCACCAACTGGGCCAGTATCGGCCAGCTCTGGGAGCTTCAAGCCTGATGACATTAACC 142
DB 84 TCACCAACTGGGCCAGTATCGGCCAGCTCTGGGAGCTTCAAGCCTGATGACATTAACC 143
QY 143 CCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 202
DB 144 CCAGCCTTTGGCACCACCTCATCTACGCTTTCGCTGGGCTGATGACCAACCCAGCTGAGCA 203
QY 203 CCATAGAAATGGAATGATCTTCTATTAAGCTTTCAATGACCTTGAAGAACAGGAAACA 262
DB 204 CCAGTGTGGAATGACAGAGACTCTCTACAGGAGTTCATGGCTGAGAGATGATC 263
QY 263 GCAAACTGAAACCTCTCTGCAATTTGAGGCTGGAATTTGGAATCTCTCTTTCACTA 322
DB 264 CCAAGCTGAAGACCTGTTAGCCATCGGAGGCTGGAATTTGGCACTCAGAACTTCACAG 323
QY 323 CCATGGTTTCCACTTCTCAGAACCCGAGACCTTCATTAACCTCAGTCAATCAATTTTCG 382
DB 324 ATATGGTAGCCGACGACCAACCTGCTGAGACTTTGTCACTCGGCCATCAGTTTCTGCG 383
QY 383 GTCAGTATGGTTTGTGAGTACCTGAGCTGGAATACCCAGCTCAGCTGGGAGCC 442
DB 384 GCAATACAGCTTTGACGGCTTGACCTTGAGTGGAGTACCCAGAGAGCAGGGAGCC 443
QY 443 CTCTCAGGACAGCATCTTTCACTGCTCTGCTGGTGAAGAAATGCGTGAAGCTTTTGAGC 502
DB 444 CTGCGGTAGACAAGAGGCTTCAACACCTGTTACAGGACTTGGCCAAATGCTTCCAGC 503
QY 503 AGGAGGCTATTGAGAGCAACAGGCCAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTG 562
DB 504 AGGAAGCCAGACCTCAGGAGGAGAACCGCTTCTTCTGAGTGCAGCGGTTCACGCTGGGC 563
QY 563 TTTCACATCCAGCTGGCTGATGATCCCTGATGATTTCTAAGTACCTGGATTTTCATCC 622
DB 564 AGACCTATGAGTCTGATGATGAGTGGACAAATCGCCAGAACCTGGATTTTGTC 623
QY 623 ATGTCATGACATATGACCTCCATGGCTCTCTGGAGGGTACACTGGGAGAAATATCTCTC 682
DB 624 ACCTATGGCTTACGACTTCCATGGCTCTTGGGAGAGGTCAAGGACATCAACAGCCCC 683
QY 683 TTTCAAAATACCTTACTGAGACTGATGCAATGCTTACCTCAATGATGATGTCATGA 742
DB 684 TCTACAGAGGCAAGAGAGTGGTGCAGCAGCCCTCAACGTGGATGCTGCTGTCG 743
QY 743 ACTATTGGAAGAACCTGAGCCAGCTGAGAGCTCATTTGTTGATTTCCAGAGATG 802
DB 744 AACAGTGGCTGCAGAGAGGGGCCCTGCGCAGAGCTGATCTTGGCATGCTTACCTAGC 803
QY 803 GACACACCTTATCTCTGAGAAACCCCTCTGTATAATGGAATTTGGTGGCCCTACCTCTG 862

DB 804 GACGCTCCTTACACTGGCCCTCTCATCAGACACCAGAGTGGGGCCCCCAGCCACAGGT 863
QY 863 ATGGCCCTGCTGGGCCCTTATACGACAGGCTGGTCTTCTGGCCCTACTATGAGATTTGCA 922
DB 864 CTGGCAGCTCCAGGCCCTTTCACCAAGGAGGAGGATGCTGGCCCTACTATGAAGTCTGCT 923
QY 923 CCTTTCTGAGAAGTGGAGCCACTGAGGTCTGGGATGCTCCCAAGAGTGGCCCTATGCT 982
DB 924 CCT-----GGAAGGGGGCCACCAACAGAGAAATCCAGATCAGAAAGTGGCCCTACATCT 977
QY 983 ATAGGCCCAAGAGTGGCTTGGCTATGACATATCAAGAGCTTCAAGTCTTAAAGGCTCAGT 1042
DB 978 TCCGGGACACCACTGCTGGGCTTTGATGATGTGGAGAGCTTCAAAAACCAAGGTCAGCT 1037
QY 1043 GCGTTAAGCAGAACAAATTTTGGAGGTGCCATGATCTGGGCCATTGACCTTGATGACTTCA 1102
DB 1038 ATCTGAACAGAAAGGAGTGGCGGGGCCATGCTGGGCACTGGCACTAGATGACTTTG 1097
QY 1103 CTGGCTCTTTCTGTGATCAGGGAATAATTTCTCTCTGA 1138
DB 1098 CCGGCTTCTCTGTGCAACCAAGGCGCGCATACCCCTCA 1133
RESULT 9
AAT50833
ID AAT50833 standard; cDNA; 1643 BP.
XX AAT50833;
AC AAT50833;
XX
XX 24-MAR-1997 (first entry)
XX Human chitinase cDNA clone chi.50.
XX
XX Chitinase; chitotriosidase; chitin; infectious disease;
KW gene therapy; vaccine; lysosomal lipodosis; Gaucher's disease;
KW leishmaniasis; sarcoidosis; X-linked adrenoleukodystrophy;
KW multiple sclerosis; drug delivery; cosmetics; food; ss.
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH CDS 13..1413
FT /*tag= a
FT sig_peptide 13..75
FT /*tag= b
FT mat_peptide 76..1410
FT /*tag= c
XX
XX WO9640940-A2.
XX
XX 19-DEC-1996.
XX
XX 06-JUN-1996; 96WO-NL00225.
XX
XX 07-JUN-1995; 95US-0486839.
XX
XX (UNAM) UNIV AMSTERDAM.
XX
XX Aerts JMF;
XX
XX WPI; 1997-118698/11.
XX
XX P-PSDB; AAW08584.
XX
XX New human chitinase - used to treat or prevent infection by
PT chitin-contg. pathogens, in diagnosis and as additives to cosmetics,
PT foods, implants etc.
XX
XX Claim 2; Page 39-40; 58pp; English.
XX
XX A cDNA clone (AAT50833), designated chi.50, codes for a 50 kDa human
CC chitinase (AAW08584) that is stable to many proteases, active at pH 3-
CC 8 and up to 50 deg, and stable in the circulation. Clones chi.50

CC and chi.39 (see also AAT50834) were isolated from a human macrophage
CC cDNA library using as probe a partial clone obtd. using primers
CC (see also AAT50835-36) based on a chitotriosidase purified from a
CC type 1 Gaucher disease patient. The 2 cDNA clones are the result
CC of alternative splicing of RNA. Chitinase nucleic acid can be used
CC for large-scale prodn. of recombinant human chitinases, or can be
CC incorporated into a gene therapy vector to treat or prevent
CC infection by chitin-contg. pathogens.
XX
SQ Sequence 1643 BP; 364 A; 490 C; 442 G; 347 T; 0 other;

Query Match 28.0%; Score 426.8; DB 18; Length 1643;
Best Local Similarity 62.1%; Pred. No. 3.4e-119;
Matches 693; Conservative 0; Mismatches 417; Indels 6; Gaps 1;

QY 23 CAGGCTGGCTCTTCTGCTGAATGCTCAGCTGGGCTCGCTGATATGCTGATATGCTAT 82
DB 35 CAGGTTTCATGCTGCTGCTGATGATCCCATGGGCTCTGCTCAAAACTGGTCTGCTACT 94
QY 83 TCACCAACTGGGCCAGTATCGGCCAGGCTCTGGGGAGCTTCAAGCCTGATGACATTAACC 142
DB 95 TCACCAACTGGGCCAGTATCGGCCAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 154
QY 143 CTGCGCTGTGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 202
DB 155 CCAGCCCTTTGCAACCCACCTCATCTAGCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 214
QY 203 CCATGAATGGAATGATGTTACTCTCTATTAAGCTTTCATGATGCTGATGATGATGATG 262
DB 215 CCAGTGGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 274
QY 263 GCAAACTGAAACCTCTCTGGAATTTGGAGCTTGGAACTTTGGAATGCTCTCTTCACTA 322
DB 275 CCAAGCTGAAGACCTGTTAGCCATCGGAGCTTGGATTTGCGCACTCAGAGTTTCACAG 334
QY 323 CCATGTTTCCACTTCTAGAACCCGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 382
DB 335 ATATGCTGACCCAGGCAACCCGCTGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 394
QY 383 GTCAGTATGGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 442
DB 395 GCATATAGCTTTGACGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 454
QY 443 CTCCTCAGACAGCATCTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 502
DB 455 CTGCGCTAGACAGGAGCTTTCACAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 514
QY 503 AGGAGCTATTGAGACACAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 562
DB 515 AGGAAGCCAGACCTCAGGGAAGGAGGCTTCTTCTGAGTGCAGCGGTTCCAGCTGGGC 574
QY 563 TTTCCAAACATCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 622
DB 575 AGACCTATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 634
QY 623 ATGTCATGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 682
DB 635 ACCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 694
QY 683 TTTACAAATACCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 742
DB 695 TCTAAGAGGCAAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 754
QY 743 ACTATTGGAAGAACATGAGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 802
DB 755 ACAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 814
QY 803 GACACCTTCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 862
DB 815 GAGCTCTCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 874
QY 863 ATGGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 922

DB 875 CTGGCACTCCAGGCCCTTCCACCAAGAGAGGAGTGTGGCTCTACTATGAAGTGTGCT 934
QY 923 CCTTTCTGAGAGTGGAGCCACTGAGTCTGGGATGCTCTCCCAAGAGTGCCTATGCCT 982
DB 935 CCT-----GGAAGGGGGCCCAACACAGAGATCCAGGATCAGAAGGTGCTCATCT 988
QY 983 ATAAGGCCAACGAGTGGCTTGGCTATGATCAATATCAAGAGCTTCAAGTGTAAAGGCTCAGT 1042
DB 989 TCCGGGACACACAGTGGTGGGCTTTGATGATGTGGAGAGCTTCAAAACCAAGTCACT 1048
QY 1043 GGTAAAGCAGACAAATTTTGGAGTGCATGATCTGGGCCATGACCTTGTGATGATCTCA 1102
DB 1049 ATCTAAGCAGAGGAGTGGGCGGGCCATGCTCTGGGCACTGGGACTTAGTACTTTG 1108
QY 1103 CTGGCTCTTCTGCTGATCAGGGAATTTCCCTCTGA 1138
DB 1109 CCGGCTCTCTCTGCAACAGGCGCGGATACCCCTCA 1144

RESULT 10
AAT50834
ID AAT50834 standard; cDNA; 1713 BP.
XX
XX AAT50834;
XX
XX 24-MAR-1997 (first entry)
XX Human chitinase cDNA clone chi.39.
XX
XX Chitinase; chitotriosidase; chitin; infectious disease;
KW gene therapy; vaccine; lysosomal lipodosis; Gaucher's disease;
KW leishmaniasis; sarcoidosis; X-linked adrenoleukodystrophy;
KW multiple sclerosis; drug delivery; cosmetics; food; ss.
XX
OS Homo sapiens.
XX
FH Key
FT CDS
FT Location/Qualifiers
FT 13..1176
FT /*tag= a
FT sig_peptide 13..75
FT /*tag= b
FT mat_peptide 76..1173
FT /*tag= c
XX
XX WO9640940-A2.
XX
XX 19-DEC-1996.
XX
XX 06-JUN-1996; 96WO-NL00225.
XX
XX 07-JUN-1995; 95US-0486839.
XX
XX (UNAM) UNIV AMSTERDAM.
XX
XX Aerts JMFG;
XX
XX WPI: 1997-118698/11.
XX
XX P-PSDB; AAW08585.
XX
XX New human chitinase - used to treat or prevent infection by
PT chitin-contg. pathogens, in diagnosis and as additives to cosmetics,
PT foods, implants etc.
XX
XX Claim 2; Page 42-43; 58pp; English.
PS
XX A cDNA clone (AAT50834), designated chi.39, codes for a 39 kDa human
CC chitinase (AAW08584) that is stable to many proteases, active at pH 3-
CC 8 and up to 50 deg, and stable in the circulation. Clones chi.39
CC and chi.50 (see also AAT50833) were isolated from a human macrophage
CC cDNA library using as probe a partial clone obtd. using primers
CC (see also AAT50835-36) based on a chitotriosidase purified from a
CC type 1 Gaucher disease patient. The 2 cDNA clones are the result
CC of alternative splicing of RNA. Chitinase nucleic acid can be used

[illegible]

SQ Sequence 1768 BP; 401 A; 529 C; 474 G; 363 T; 1 other;
SQ Sequence 1768 BP; 401 A; 529 C; 474 G; 363 T; 1 other;

Sequ

Db	1160	ATCTGAAGCAGAAAGGACTGGCGGGGCCATGGTCTGGGCACTTAGACTTACACTTTG	1219
Oy	1103	CTGGCTCTTCTGTGTGATCAGGGAATAATTTCTCTGA	1138
Db	1220	CCGCTTCTCTGCAACACGAGCGGATACCCCTCA	1255
RESULT 12			
AAV10436			
ID	AAV10436	standard; cDNA; 1656 BP.	
XX	AAV10436;		
XX	15-JUN-1998	(first entry)	
XX	Human chitinase clone MO-13B cDNA.		
XX	Chitinase; human; fungal infection; immunogen; diagnosis; treatment;		
KW	Gaucher's disease; transgenic; detection; hybridisation; antifungal;		
KW	rheumatoid arthritis; overexpression; extracellular matrix; ss.		
OS	Homo sapiens.		
XX	Key	Location/Qualifiers	
FH	CDS	27..1427	
FT		/*tag= a	
FT	sig_peptide	27..89	
FT		/*tag= b	
FT	mat_peptide	90..1424	
FT		/*tag= c	
FT		/product= chitinase	
FT		/note= "from clone MO-13B"	
XX	WO9747752-A1.		
PN	18-DEC-1997.		
XX	16-JUN-1997;	97WO-US10460.	
XX	14-JUN-1996;	96US-0663618.	
XX	(ICOS-) ICOS CORP.		
XX	Gray PW;		
XX	WPI; 1998-052316/05.		
DR	P-PSDB; AAW40260.		
XX	Nucleic acids encoding human chitinase - useful as antifungal		
PT	agents, especially in combination with other antifungals		
XX	Claim 9; Page 42-44; 63pp; English.		
PS	This sequence encodes a novel human chitinase isolated from clone MO-13B.		
XX	Chitinases are useful for treating or preventing fungal infection and		
CC	as immunogens for generating antibodies which are used to purify, detect		
CC	and quantify chitinases, e.g. for diagnosis of Gaucher's disease. The		
CC	nucleic acid sequence of the chitinase is also useful as a probe to		
CC	identify and isolate genomic DNA encoding chitinases or similar proteins.		
CC	or cells expressing them or to generate transgenic ('knockout') rodents.		
CC	It can also be used in hybridisation assays and to detect genetic		
CC	alterations in the chitinase gene related to disease. Agents that inhibit		
CC	this protein may be useful in treatment of Gaucher's disease and		
CC	rheumatoid arthritis, where overexpression of the protein can damage		
CC	the extracellular matrix. Chitinase also improves the activity of other		
CC	antifungal agents and may allow a reduction in the dose of such agents,		
CC	and thus of their side effects.		
XX	Sequence 1656 BP; 365 A; 497 C; 447 G; 347 T; 0 other;		
XX	Query Match	27.9%; Score 425.2; DB 19; Length 1656;	
XX	Best Local Similarity	62.0%; Pred. No. 1.1e-118;	

Query Match 27.9%; Score 425.2; DB 19; Length 1656;
Best Local Similarity 62.0%; Pred. No. 1.1e-118;

Qy	1103	CTGGCTCTTTCTGTGATCAGGAAAAATTTCTCTGA	1138
Db	1123	CGGGCTTCTCTGCAACCAGGCGGATACCCCTCA	1158
RESULT 13			
AA	AAZ21848		
ID	AAZ21848	standard; DNA; 1656 BP.	
XX	AAZ21848;		
AC			
XX	10-DEC-1999	(first entry)	
XX			
DE	MO-13B	clone of human Chitinase, with noncoding 5'/3' regions.	
XX			
DE	chitin; fungal infection; immunocompromised; AIDS; chemotherapy;		
KW	organ transplant; parasite; chitin-binding; allele; vector;		
KW	truncated protein; ds.		
XX			
XX	Homo sapiens.		
OS			
XX			
FH	Key	Location/Qualifiers	
FT	CDS	27..1427	
FT		/*tag= a	
FT		/product= Human_Chitinase	
FT	sig_peptide	27..89	
FT		/*tag= b	
FT		/note= "Signal peptide"	
FT	mat_peptide	90..1427	
FT		/*tag= c	
FT		/note= "Mature peptide"	
XX			
XX	W09946390-Al.		
PN			
XX			
XX	16-SEP-1999.		
PD			
XX			
XX	12-MAR-1999;	99WO-US05343.	
XX			
XX	12-MAR-1998;	98US-0039198.	
PR			
XX			
XX	(ICOS-) ICOS CORP.		
PA			
XX			
PI	Gray PW, Tjoelker LW;		
XX			
XX	WPI; 1999-551417/46.		
DR	P-PSDB; AAY42426.		
XX			
PT	Novel chitin-binding fragments of human chitinase used to treat fungal		
PT	infections in animals		
XX			
XX	Example 1; Page 59-62; 83pp; English.		
PS			
XX			
CC	This is the nucleotide sequence of an allelic form of the human		
CC	chitinase enzyme, which is capable of degrading Chitin (a linear		
CC	homo polymer of beta-1,4-linked N-acetylglucosamine residues).		
CC	Chitinase fragments can be used to screen for proteins or other		
CC	molecules that specifically bind to the chitin-binding domain of human		
CC	chitinase or that modulate its activity. These compounds are useful for		
CC	immunization, as well as for purifying chitinase, as well as for		
CC	detection and quantification of chitinase. Polynucleotide fragments of		
CC	the invention are useful as a source of probes and primers, and to		
CC	express the proteins recombinantly. The chitinase fragments, when		
CC	conjugated to antifungal compounds, are used to treat animals,		
CC	especially humans, infected with chitin-containing parasites such as		
CC	fungi. Fungal infection treated include candidiasis, aspergillosis,		
CC	coccidioidomycosis, blastomycosis, paracoccidioidomycosis,		
CC	mucomycosis, histoplasmosis, cryptococcosis, chromoblastomycosis,		
CC	sporotrichosis, and dermatophytoses.		
CC	Chitin can be degraded by the enzyme chitinase. Use of whole chitinase		
CC	protein for treating infections, especially fungal infections, is		
CC	problematic. In view of the increasing incidents of life- threatening		
CC	fungal infection in e.g. immunocompromised individuals, there exists a		

CC methods for the recombinant production of human chitinase products,
CC which are expected to be useful as products for treating fungal
CC infections or for the development of such products. Human
CC chitinase has a synergistic effect on the actions of other
CC fungicides. It can be administered to improve the antifungal
CC activity of a non-chitinase antifungal agent, especially
CC amphotericin B or itraconazole, in the treatment of a fungal
CC infection such as candidiasis, aspergillosis, coccidioidomycosis,
CC blastomycosis, paracoccidioidomycosis, histoplasmosis,
CC cryptococcosis, chromoblastomycosis, sporotrichosis, mucormycosis,
CC dermatophytoses and Pneumocystis infections (all claimed). In
CC particular, the fungal infection involves Candida, Aspergillus
CC and/or Cryptococcus spp.; whose growth is not effectively
CC inhibited by contact with human chitinase alone.

xx
SQ Sequence 1656 BP; 365 A; 497 C; 447 G; 347 T; 0 other;

Query Match 27.9%; Score 425.2; DB 24; Length 1656;
Best Local Similarity 62.0%; Pred. No. 1.1e-118;
Matches 692; Conservative 0; Mismatches 418; Indels 6; Gaps 1;
QY 23 CAGTCTGGCTCTTCTGCTGAATGCTCAGCTGGGCTGCTACAATCTGATATGCTATT 82
DB 49 CAGTTTCATGGTCTGCTGATGATCCATGGGGCTCTGCTGCAAACTGGTCTGCTACT 108
QY 83 TCACCAACTGGGCCCATATCGCCAGGTCTGGGAGCTTCAAGCCTGATGACATTAAAC 142
DB 109 TCACCAACTGGGCCCATATCGCCAGGTCTGGGAGGCTGCTTCTGCTGCCAAGGACTGGACC 168
QY 143 CTTGCTGCTGATCACCCTGATCTATGCTTCTGCTGGGATGAGACATGATCAACA 202
DB 169 CCAGCCTTTGACCCCACTCATCTAGCCTTCTGCTGATGACCAACACCACTGAGCA 228
QY 203 CCATAGATGAATGATGTTTACTCTCTATAAGCTTTCAATGACTTGAATAACAGGAACA 262
DB 229 CCAGTGAATGATGAGGACTCTCTACAGAGTTCAATGGCTTGAAGAATGAATC 288
QY 263 GCAAACTGAAACCCCTCTGGCAATTTGGAGCTTGAACCTTTGGAAGTCTCTTTCACTA 322
DB 289 CCAAGCTGAAGACCCCTGTTAGCCATCGGAGCTTGAATTTTCAAGCTCAGAACTTCACAG 348
QY 323 CCATGTTTCCACTTCTCAGAACCCGACAGCTTCAATACCTCAGTCAATCAAAATTTCTGC 382
DB 349 ATATGATAGCCGCGCAACACCGTCAGACCTTTGTCACTCGGCATCAGGTTTCTGC 408
QY 383 GTCAGTATGGTGTGATGAGTGGACTGGACTGGGAATACCAGGCTCACGTGGGAGCC 442
DB 409 GCAATACAGCTTTGACGGCTTTGACCTTGACTGGGAGTACCCAGGAAGCCAGGGAGCC 468
QY 443 CTCCTCAGGACAAGCATCTTCTCAGTCTGCTGGTGAAGGAATCGGTGAAGCTTTTGAGC 502
DB 469 CTGCCGTAGACAAGGAGCGCTTCAACACCTGCTACAGGACTTGGCCAATGCCCTTCCAGC 528
QY 503 AGGAGCTATTGAGACAGACAGCCGACAGTATGTTACTGCTGTGATGCTGGTGGGA 562
DB 529 AGGAAGCCCAAGACCTCAGGGAAGAACCGCTTCTTCTGAGTGAGCGGTTCCAGCTGGGC 588
QY 563 TTTTCAACATCCAGGCTGGCTATGATCCCTGAACCTTTCTAAGTACCTGGATTTTCATCC 622
DB 589 AGACCTATGTGGATGCTGGATAGGATGGGACAAATCGCCCAAGACCTTGGATTTGTCA 648
QY 623 ATGTCATGACATATGACCTCCATGGCTTCTGGAGGGCTACACTGGGAGGAATAGTCCCTC 682
DB 649 ACCTTATGGCTACGACTTCCATGGCTTGGGAGAGGTCACGGGACATAACAGCCCC 708
QY 683 TTTTACAATACCTACTGAGCTGGTAGCAATGCCCTACCTCAATGCTGGATTTGTCATGA 742
DB 709 TCTACAAGAGGCAAGAGAGTGGTGGACAGCCAGCCTCAACGCTGGATGCTGTGTC 768
QY 743 ACTATTGGAGAAACAATGAGCCCAAGCTGAGAGCTCATTTGTTGATTCCCAAGATATG 802
DB 769 AACAGTGGCTGCAGAGGGGACCCCTGCGCAGCAAGCTGATCTCTTGGATGCCCTACCTACG 828

Search completed: July 3, 2003, 05:03:20
Job time : 351.54 secs

QY 803 GACACACCTTCTATCTGTGAGAAACCCCTCTGATTAATGAAATGTGCTCCCTACCTCTGGTG 862
DB 829 GAGCTCTCTTACACTGGGCTCTCTCATCAGACACAGAGTGGGGGCCCCAGCCACAGGT 888
QY 863 ATGGCCCTGCTGGCGCTTATACCAGACAGCTGGGTTCTGGGCTTACTATGAGATTGCA 922
DB 889 CTGGCACTCCAGGCCCTTTCACCAAGGAAGGAGGATGCTGGCTTACTATGAACTCTGCT 948
QY 923 CTTTCTGAGAAGTGGAGCCACTGAGGTCTGGGATGCTCTCCCAAGAAGTGCCTATGCCCT 982
DB 949 CCT-----GGAAGGGGGCCCAACACAGAGATCCAGGATCAGAAGTGCCTTACATCT 1002
QY 983 ATAAGCCCAACAGTGGCTTGGTATGATCAATATCAAGAGCTTCACTGTTAAGGCTCAGT 1042
DB 1003 TCCGGGACAACCACTGGTGGGCTTTGATGATCTGGAGAGCTTCAAAAACCAAGGTCAGCT 1062
QY 1043 GGCTTAAGCAGAACAAATTTGGAGGTGCCATGATCTGGCCATTGACCTTGATGACTTCA 1102
DB 1063 ATCTGAAGCAGAAAGGAGCTGGGGGGCCCATGCTCTGGGCACTGGACTTAGATGACTTTG 1122
QY 1103 CTGGCTCTTCTGCTGATCAGGGAATAATTTCTCTCTGA 1138
DB 1123 CCGGCTTCTCTGCAACCAAGGGCCGATACCCCTCA 1158

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2003, 04:38:16 ; Search time 73.5873 seconds
(without alignments)
6355.472 Million cell updates/sec

Title: US-10-004-219B-3

Perfect score: 1525

Sequence: 1 atggccaagctactctcgt.....taaaattgtagcacaaca 1525

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA.*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PCrUS_COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	426.8	28.0	1636	US-09-039-198A-1	Sequence 1, Appl
2	426.8	28.0	1636	US-08-877-599-1	Sequence 1, Appl
3	426.8	28.0	1636	US-09-267-574-1	Sequence 1, Appl
4	426.8	28.0	1643	US-08-486-839-3	Sequence 3, Appl
5	426.8	28.0	1643	US-09-151-011-3	Sequence 3, Appl
6	426.8	28.0	1643	US-09-343-623-3	Sequence 3, Appl
7	426.8	28.0	1713	US-08-486-839-5	Sequence 5, Appl
8	426.8	28.0	1713	US-09-151-011-5	Sequence 5, Appl
9	426.8	28.0	1713	US-09-343-623-5	Sequence 5, Appl
10	425.2	27.9	1656	US-09-039-198A-3	Sequence 3, Appl
11	425.2	27.9	1656	US-08-877-599-3	Sequence 3, Appl
12	425.2	27.9	1656	US-09-267-574-3	Sequence 3, Appl
13	343.2	22.5	1433	US-08-694-915-1	Sequence 1, Appl
14	330.8	21.7	1526	US-08-694-915-3	Sequence 3, Appl
15	314.8	20.6	1681	PCT-US94-07754-4	Sequence 4, Appl
16	311.6	20.4	1681	US-08-581-527-4	Sequence 4, Appl
17	140	9.2	2452	US-08-524-051-1	Sequence 1, Appl
18	126	8.3	1677	US-09-545-814-13	Sequence 13, Appl
c 19	126	8.3	1677	US-09-545-814-15	Sequence 15, Appl
c 20	126	8.3	1749	US-09-545-814-4	Sequence 4, Appl
c 21	126	8.3	1749	US-09-545-814-6	Sequence 6, Appl
c 22	126	8.3	2610	US-09-545-814-1	Sequence 1, Appl
c 23	126	8.3	2610	US-09-545-814-3	Sequence 3, Appl
c 24	124.4	8.2	1919	US-09-545-814-31	Sequence 31, Appl
c 25	124.4	8.2	1919	US-09-545-814-33	Sequence 33, Appl
c 26	121.4	8.0	1608	US-09-292-225-20	Sequence 20, Appl
c 27	121.4	8.0	1608	US-09-292-225-22	Sequence 22, Appl

28	121.4	8.0	1665	4	US-09-292-225-17	Sequence 17, Appl
c 29	121.4	8.0	1665	4	US-09-292-225-19	Sequence 19, Appl
c 30	121.4	8.0	1752	4	US-09-292-225-14	Sequence 14, Appl
c 31	121.4	8.0	1752	4	US-09-292-225-16	Sequence 16, Appl
c 32	112.8	7.4	1470	4	US-09-292-225-40	Sequence 40, Appl
c 33	112.8	7.4	1470	4	US-09-292-225-42	Sequence 42, Appl
c 34	112.8	7.4	1527	4	US-09-292-225-37	Sequence 37, Appl
c 35	112.8	7.4	1527	4	US-09-292-225-39	Sequence 39, Appl
c 36	112.8	7.4	1621	4	US-09-292-225-34	Sequence 34, Appl
c 37	112.8	7.4	1621	4	US-09-292-225-36	Sequence 36, Appl
c 38	109.6	7.2	1478	4	US-09-545-814-28	Sequence 28, Appl
c 39	109.6	7.2	1478	4	US-09-545-814-30	Sequence 30, Appl
40	64	4.2	147	4	US-09-156-856-8	Sequence 8, Appl
41	54.8	3.6	1167	1	US-07-939-501A-6	Sequence 10, Appl
42	54.8	3.6	1167	4	US-08-448-398-10	Sequence 6, Appl
43	54.8	3.6	1320	1	US-07-939-501A-14	Sequence 14, Appl
44	54.8	3.6	1364	1	US-07-939-501A-13	Sequence 13, Appl
45	54.8	3.6	1405	1	US-07-939-501A-9	Sequence 9, Appl

ALIGNMENTS

RESULT 1
US-09-039-198A-1
; Sequence 1, Application US/09039198A
; Patent No. 6200951
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; APPLICANT: Tjoelker, Larry W.
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive/6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/039,198A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 27866/34391
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1636 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..1399
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 65..1399
US-09-039-198A-1

Query Match 28.0%; Score 426.8; DB 4; Length 1636;
Best Local Similarity 62.1%; Pred. No. 6.6e-123;
Matches 693; Conservative 0; Mismatches 417; Indels 6; Gaps 1;

QY	23	CAGGTCCTGGCTCTTCTGCTGAATGCTCAGCTGGGCTCGCCTCAACAATCTGATGATGCTATT	82
DB	24		83
QY	83	TCACCAACTGGCCCAAGTATCGGCCAGGCTCTGGGAGCTTCAAGCCTGATGACATTAACC	142
DB	84	TCACCAACTGGGCCCAAGTACAGACAGGGGAGGCTCGCTTCTGCCAAGAGCTTTGGACC	143
QY	143	CCTGCCTGTGTACTCACTGATCTATGCTTTGTGGGATGCAGAACAAATGAGATCAACA	202
DB	144		203
QY	203	CCATAGTAATGGAATGATGTTACTCTCTATAAAGCTTTCAATGACTTGAAGAACAGGAACA	262
DB	204	CCACTGAGTGAATGACGAGACTCTCTACCAAGAGTTCAATGGCTGTAAGAAGATGAATC	263
QY	263	GAAACTGAAAACCCCTCTCTGCAANTGGAGGCTGGAAGCTTTGGAACTGCTCTCTTCACTA	322
DB	264	CCAAGCTGAAACCCCTGTTAGCCATCGGAGGCTGGAATTTGGGCACTCAGAAGTTTCAAG	323
QY	323	CCATGGTTTCCACTTCTCAGAACCGCCAGACCTTCAATTAACCTCAGTCAATCAAAATTTCTGC	382
DB	324	ATATGGTAGCCACGGCCAAACACGTCAGACCTTTGTCAACTCGGCCATCAGTTTCTTGC	383
QY	383	GTCAGTATGGTTTTGATGGACTGACCTGGACTGGGAATACCCAGGCTCAGTGGGAGCC	442
DB	384	GCAAAATACAGCTTTGAGGGCCCTTGACCTTGACTGGGAGTACCCAGGAAGCCAGGGGAGCC	443
QY	443	CTCCCTCAGGACAAGCATCTCTTCACTGTCTCTGGTGAAGGAATGCGTGAAGCTTTTGAGC	502
DB	444	CTGCCGTAGACAAGGAGCGCTTCAACACCTGGTACAGGACTTGGCCATGCCCTTCCAGC	503
QY	503	AGGAGGCTATTGAGAGCAACAGGCCCAAGACTGATGGTTACTGCTGCTGTAGCTGGTGGGA	562
DB	504	AGGAAGCCCAAGACCTCAGGGAAGAACGCTTCTTCTGAGTGCAGCGGTTCCAGCTGGC	563
QY	563	TTTCCAACATCCAGGCTGGCTATGAGATCCCTGAACTTCTTAAGTACCTGGAATTTCACTCC	622
DB	564	AGACCTATGTGGATGCTGTGATACAGAGTGGACAAATTCGCCCAAGAACTTGAATTTGTCA	623
QY	623	ATGTCATGACATATGACCTCATGGCTCTGGGAGGGCTACACTGGGGAGAATAGTCTCTC	682
DB	624	ACCTTATGGCTTAGACTTCCATGGCTCTTGGGAGAAGTACACGGGACATACAGACCCCC	683
QY	683	TTTACAATACCTTACTGAGACTGGTAGCAATGCCTACCTCAATGTGSAATATGATCATGA	742
DB	684	TCTACAAGGCGCAAGAAGATGGTGCAGCAGCCAGCCTCAACGTGGATGCTGCTGTGC	743
QY	743	ACTATTGGAGAACAAATGGAGCCCACTGAGAGCTCATTTGTTGGATTCCCACAGTATG	802
DB	744	AACAGTGGCTGCAAGAGGGACCCCTGCCAGCAAGCTGATCCTTTGGCATGCCCTACCTACG	803
QY	803	GACACACTTTCATCCTTGAGAAACCCCTCTGATATGGAATTGGTGCCTTACCTTCTGGTG	862
DB	804	GAGGCTCTTCACACTGSCCTCTCATCAGACACCAAGTGGGGGCCCCCAGCACAGGTT	863
QY	863	ATGGCCCTGCTGGGCCCTATACACAGACGGCTGGGTTCTGGGCCCTACTATGAGATTTGCA	922
DB	864	CTGGCACTCAGGCCCTTCCACCAAGGAAGAGGGGATGCTGGCCCTACTATGAAGCTGTCT	923
QY	923	CCTTTCTGAAAGTGGAGCCACTGAGTCTGGGATGCCCTCCCAAGAAAGTGCCTATGCCCT	982
DB	924	CC1-----GGAAGGGGGCCACCAACAGAGAAATCCAGGATFCAGAAGGTGCCCTTACATCT	977
QY	983	ATAAGGCCAACGAGTGGCTTGGGTATGACAATATCAAGAGCTTTCAGTGTTAAGGCTCAGT	1042
DB	978	TCGGGGACACCAAGTGGTGGCTTTTGATGATGTGGAGAGCTTCAAAACCAAGGTCAGCT	1037
QY	1043	GGCTTAAGCAGACAAATTTTGGAGTGCCCATGATCTGGGCCAATGACCTTGATGACTTCA	1102
DB	1038	ATCTGAGCAGAGAGGACTGGGGGGGCCATGGCTGGGCACGTGACTTATGATGACTTTG	1097

Qy 1103 CTGGCTCTTCTGTGATCAGGGAAAATTTCCCTCTGA 1138
| | | | | | | | | | | | | | | | | | | | | |
db 1098 CCGGCTTCTCCTGCAACCAGGGCCGATACCCCTCA 1133

RESULT 2

US-08-877-599-1
: Sequence 1, Application US/08877599
: Patent No. 6372212
: GENERAL INFORMATION:
: APPLICANT: Gray, Patrick W.
: TITLE OF INVENTION: Chitinase Materials and Methods
: NUMBER OF SEQUENCES: 17
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
: STREET: 6300 Sears Tower, 233 South Wacker Drive
: CITY: Chicago
: STATE: Illinois
: COUNTRY: United States of America
: ZIP: 60606-8402
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/877,599
: FILING DATE:
: CLASSIFICATION: 514
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/663,618
: FILING DATE: 14-JUN-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Rin-Laures, Li-Hsien
: REGISTRATION NUMBER: 33,547
: REFERENCE/DOCKET NUMBER: 27866/33994
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 312/474-6300
: TELEFAX: 312/474-0448
: TELEX: 25-3856
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1636 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 2..1399
: FEATURE:
: NAME/KEY: mat_peptide
: LOCATION: 65..1399
: US-08-877-599-1

Query Match	28.0%;	Score 426.8;	DB 4;	Length 1636;
Best Local Similarity	62.1%;	Pred. No. 6.6e-123;		
Matches 693; Conservative	0;	Mismatches 417;	Indels 6;	Gaps 1;

Qy	23	CAGGTC	TGGCTCTTCTGCTGAATGCTCAGCTGGGGCTGCTCAATCTGATATGCTATT	82
Db	24	CAGGTT	CAATGGTCTGCTGATGATCCATGGGCTCTGTCGCAAACTGGTCTGCTACT	83
Qy	83	TCACCA	CACTGGGCCAGATCGCCAGGCTCTGGGGAGCTTCAAGCTGTAGCATTAACC	142
Db	84	TCACCA	CACTGGGCCAGTACAGACAGGGGAGGCTCGCTTCTGCCCAAGACTTGGACC	143
Qy	143	CTGCTCT	GTGTACTCACCTGATCTATGCTTTGCTGGGATGCAGAAACAATGAGATCACCA	202
Db	144	CCAGCCT	TTGCACCAACCTCATCTACGSCCTTCGCTGGCATGACCAACACCACGCTGAGCA	203
Qy	203	CCATAG	AATGGAATGATTTTACTCTCTATAAAGCTTTCAATGACTTTGAAAAACAGGAACA	262

Db 204 CCAGTGTAGTGGAAATGACGAGACTCTCTACAGGAGTTCAATGGCTTGAAGAAGATGAATC 263
QY 263 GAAACCTGAAACCTCCCTGGCAATTTGGAGGCTGGAACTTTGGAACCTGCTCTTTCACATA 322
Db 264 CCAAGCTGAAGACCTCTGTAGCCATCGGAGGCTGGAAATTCGGCACTCAGAGTTCCACAG 323
QY 323 CCATGGTTTCCACTTCTCAGAACCCGACAGCTTTCATTACCTCAGTCATCAAAATTTCTGC 382
Db 324 ATATGGTAGCCACGCGCAACACCGTTCAGACCTTTGTCAACTCGGCCATCAGTTTCTGC 383
QY 383 GTCAATATGGTTTGTATGACTGGACCTGGACCTGGGAATACCCAGGCTCAGCTGGAGCC 442
Db 384 GAAATATACAGCTTTGACGGCTTGTACCTTGTAGTGGAGTACCCAGGAAGCCAGGGAGCC 443
QY 443 CTCCTCAGGACAGCATCTCTTACCTGCTCTGGTGAAGAAATCGTGAAGCTTTTGAGC 502
Db 444 CTGCGGTAGCAAGGAGCGCTTCACAACTTGTACAGGCTTGGCCATGCTTCCAGC 503
QY 503 AGGAGCTATTGAGAGCAACAGGCCAGACTGATGTTACTGCTGTAGCTGTGGTGGGA 562
Db 504 AGGAAGCCAGACCTCAGGGAAGGAGCGCTTCTCTGAGTGCAGCGGTTCCAGCTGGGC 563
QY 563 TTTCCAACTCAGGCTGCTATGAGATCCCTGAACTTCTTAGTACCTGGAATTCATCC 622
Db 564 AGACCTATGTGATGTGATACGAGGTGGACAAATCGCCAGAACCTGGATTTGTGA 623
QY 623 ATGTATGATACATGATGACCTCCATGGCTCTGGAGGCTTACACTGGGAGAAATAGCTC 682
Db 624 ACCTTATGGCTTACGACTTCCATGGCTCTTGGAGAGGTACCGGACATACAGCCGCC 683
QY 683 TTTACAAATACCCCTACTGAGACTGTGAGTACCTTCTAGTACCTGGAATTCATCC 742
Db 684 TCTAAGAGGCAAGAGAGTGTGTGACAGCCAGCTCAACGTGGATGCTGTGTGC 743
QY 743 ACTATTGGAGAACATGAGCCCGCAGCTGAGAGCTCATTTGATTTCCAGAGTATG 802
Db 744 AACAGTGGCTGCAGAGGGGAGCCCTGCGCAGCAAGCTGATCTTGGCATGGCTTACACG 803
QY 803 GACACACTTCTCTGAGAAACCCCTCTGATAATGGAATTTGGTGGCCCTACTCTGTGTG 862
Db 804 GAGCTCTCTACACTGGCTCTCTATCAGACACAGAGTGTGGGCGCCAGCCAGAGGT 863
QY 863 ATGGCCCTGCTGGGCTATACAGACAGGCTGGTGTCTGGGCTTACTATGAGATTGCA 922
Db 864 CTGGCACTCCAGGCCCTTCCACCAAGGAAGAGGATGCTGGCTTACTATGAATCTGCT 923
QY 923 CTTTCTGAGAGTGGAGCCACTGAGTCTGGATGCCCTCCCAAGAGTGGCCCTATGCT 982
Db 924 CCT-----GGAGGGGGCCCAACAGAGAAATCCAGATCAGAGAGTGGCCCTACATCT 977
QY 983 ATAGGCCAACGAGTGGCTTGGCTATGACAATATCAAGAGCTTCACTGTTAAGGCTCAGT 1042
Db 978 TCCGGACAAACAGTGGGTGGCTTGTATGATGTGGAGAGCTTCAAAACAGAGTCACT 1037
QY 1043 GGTAAAGCAGAACAAATTTGGAGTGGCCATGATCTGGGCCATTTGACCTTGTATGATCA 1102
Db 1038 ATCTGAAGCAGAGGAGTGGGCGGGCCATGCTCTGGCACTGGACTTAGATGACTTTG 1097
QY 1103 CTGGCTCTTCTGTATGATCAGGAAATTTTCTCTGA 1138
Db 1098 CCGGCTTCTCTGTCAACACGAGGGCCGATACCCCTCA 1133

RESULT 3

US-09-267-574-1
; Sequence 1, Application US/09267574
; Patent No. 6399571
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; APPLICANT: Tjoelker, Larry W.
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
; FILE REFERENCE: 27866/35407
; CURRENT APPLICATION NUMBER: US/09/267,574

; CURRENT FILING DATE: 1999-03-12
; EARLIER APPLICATION NUMBER: 09/039,198
; EARLIER FILING DATE: 1998-03-12
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1636
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(1399)
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (65)..(1399)
US-09-267-574-1

Query Match 28.0%; Score 426.8; DB 4; Length 1636;
Best Local Similarity 62.1%; Pred. No. 6.6e-123;
Matches 693; Conservative 0; Mismatches 417; Indels 6; Gaps 1;

QY 23 CAGGCTGTGCTCTTCTGCTGAATGCTCAGCTGGGGTCTGCCTACATCTGATGATGATTT 82
Db 24 CAGGTTTCATGTGCTCTGCTGATGATCCCATGGGCTCTGCTGCAAACTGGTCTGCTACT 83
QY 83 TCACCAACTGGGCGCCAGTATCGGCCAGGTCTGGGAGCTTCAAGCCTGATGACATTAACC 142
Db 84 TCACCAACTGGGCGCCAGTATCAGACAGGGGAGGCTCGCTCTGCCCAAGGACTTGGACC 143
QY 143 CTTGCTGTGTACTACCTGATGATGATGCTTGTGCTGGGATGAGAACAAATGATGATCACC 202
Db 144 CCAGCTTTGCAACCCACCTCATCTAGCCCTTCTGCTGGCATGACCAACCCAGCTGAGCA 203
QY 203 CCATAGAATGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 262
Db 204 CCACGTAGGGAATGAGAGACTCTCTACAGAGTTCATAGGCTTGAAGAAGATGAATC 263
QY 263 GCAAACTGAAACCTCTCTGGCAATTTGGAGGCTTGAACCTTTGGAACCTGCTCTTTCACATA 322
Db 264 CCAAGCTGAAGACCTCTTGTAGCCATCGGAGGCTGGAATTTCCGCACTCAGAACTTCCAG 323
QY 323 CCATGGTTTCCACTCTCAGAACCCGACAGCTTCAATACCTCAGTCATCAAAATTTCTGC 382
Db 324 ATATGTAGTACCGGCCAACACCGTCCAGACCTTTGCTCAACTCGGCCATCAGGTTTCTGC 383
QY 383 CTCAGTATGGTTTGTAGTACCTGAGCTGGAGTGGGAATACCCAGCTCAGCTGGGAGCC 442
Db 384 GCAAAATACAGCTTTGAGCGCTTGAACCTTGAAGTGGAGTACCCAGGAAGCCAGGGAGCC 443
QY 443 CTCCTCAGGACAGCATCTCTTCACTGTCTGTTGAAGGAAATCGCTGAAAGCTTTTGGAGC 502
Db 444 CTGCGCTACAAAGGAGCGCTTCAACACCTGTTACAGGACTTGGCCATGCTCTTCCAGC 503
QY 503 AGGAGCTATTGAGAGCAACAGGCCAGACTGATGTTTACTGCTGTAGTGTGGTGGGA 562
Db 504 AGGAAGCCAGACCTCAGGGAAGGAGCGCTTCTCTGAGTGCAGCGGTTCACAGCTGGGC 563
QY 563 TTTCCAACTCAGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 622
Db 564 AGACCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 623
QY 623 ATGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 682
Db 624 ACCTTATGGCTTACGACTTCCATGGCTCTTGGGAGAGGTCAGGGGACATACAGGCCGCC 683
QY 683 TTTACAAATACCCCTACTGAGACTGTGAGCAATGCCCTACCTCAATGTGGATTTATGTCATGA 742
Db 684 TCTAAGAGGCAAGAGAGTGTGTGACAGCCAGCTTCAACAGCTTCAACAGCTGCTGCTGTC 743
QY 743 ACTATTGGAGAACAAATGAGCCCGCAGCTGAGAGCTCATTTGATTTCCAGAGTATG 802
Db 744 AACAGTGGCTGCAGAGGGGAGCCCTGCGCAGCAAGCTGATCTTGGCATGGCTTACACG 803

QY 803 GACACACCTTCATCTCGAGAACCCCTCTGATGATGAATGGTGGCCCTACCTCTGGTG 862
DB 804 GACGCTCTTCACACTGGCTCTCTCATCAGACACAGAGTGGGGCCCCAGCCACAGGGT 863
QY 863 ATGGCCCTGCTGGGCCCTATACACAGACAGCTGGTCTGGGCCCTACTATGAGATTTGCA 922
DB 864 CTGGCACTCCAGGCCCTTCACCAAGGAAGAGGAGTCTGGCCCTACTATGAAAGTCTGCT 923
QY 923 CCTTCTGAGAAGTGGAGCCACTGAGTCTGGGATGCTCCCAAGAGTGGCCCTATGCT 982
DB 924 CCT-----GGAAGGGGCCCAACAACAGAGATCCAGGATCAGAGGTGCCCTACATCT 977
QY 983 ATAAGGCCAACAGTGGCTTGGCTATGACAAATACAGAGCTTCAGTGTAAAGGCTCAGT 1042
DB 978 TCCGGGACAAACAGTGGTGGCTTTGATGATGTTGGAGAGCTTCAAAAACCAAGTCACT 1037
QY 1043 GGCTTAGCAGAACAAATTTGGAGGTGCCATGATCTGGGCCATGACCTTGTAGTCTCA 1102
DB 1038 ATCTGAAGCAGAGGAGTGGGGGGCCATGCTGGGCACCTGGACTTAGATGACTTTG 1097
QY 1103 CTGGCTCTTCTGATCAGGGAATAATTTCTCTGA 1138
DB 1098 CCGGCTCTCTGACACAGGCGCGATACCCCTCA 1133

RESULT 4

US-08-486-839-3
; Sequence 3, Application US/08486839
; Patent No. 5928928
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: A human chitinase, its recombinant
; TITLE OF INVENTION: production, its use for decomposing chitin, its use
; TITLE OF INVENTION: in therapy or prophylaxis against infection diseases.
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11758
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,839
; FILING DATE: 07 - June - 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Baron, Ronald J.
; REGISTRATION NUMBER: 29,281
; REFERENCE/DOCKET NUMBER: 294-26
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1643 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; US-08-486-839-3

Query Match 28.0%; Score 426.8; DB 2; Length 1643;
Best Local Similarity 62.1%; Pred. No. 6.6e-123;
Matches 693; Conservative 0; Mismatches 417; Indels 6; Gaps 1;
QY 23 CAGGTCTGCTCTTCTGCTGAATGCTCAGCTGGGGTCTGCCTACAAATCTGATATGCTATT 82

DB 35 CAGGTCTGCTCTGCTGATGATCCATGGGGCTCTGCTCCAAACTGGTCTGCTACT 94
QY 83 TCACAACTGGGCCAGTATCGGCAGGTCTGGGAGCTTCAAGCCGTGATGACATTAACC 142
DB 95 TCACAACTGGGCCAGTACAGACAGGGGGAGTCTGCTTCTTCCCTGCCAAGAGCTTGGACC 154
QY 143 CTTGCTGCTGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 202
DB 155 CCAGCTTTCCACCCACTCTACCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 214
QY 203 CCATAGAATGAATGATGTTACTCTCTATATAAGCTTTCAATGACTTTGAAACACAGGAACA 262
DB 215 CCAGTGAATGAATGACGAGACTCTCTACAGGAGTTCATGCTTCCCTGAGCAAGATGAATC 274
QY 263 GCAAACTGAAACCCCTCTGGCAATTTGGAGGCTTGGAACTTTGGAACCTGCTCTTCACTA 322
DB 275 CCAAGCTGAAGACCTTTAGCCATCGGAGGCTTGGAAATTTCCGCACTCAGAAGTTCACAG 334
QY 323 CCATGTTTCCACTTCTCAGAACCCGACAGACTTCATTAACCTCAGTCACTCAAAATTTCTGC 382
DB 335 ATATGTTAGCCAGCCCAACACCGTCAGACCTTTGTCACTCGGCCATCAGGTTTCTGC 394
QY 383 GTCAGTATGGGTTTGTATGACTGGACCTGGACTTGGAAATACCAGGCTCACTGGGAGCC 442
DB 395 GCAATACAGCTTTGACGGCTTTGACCTTGACTGGAGTACCAGGAAAGCCAGGGAGCC 454
QY 443 CTCTCAGGACAGACTCTCTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 502
DB 455 CTGCTGACAGAGGAGGCTTCCAAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 514
QY 503 AGGAGCTATTGAGAGCAACAGCCGACAGCTGATGTTACTGCTGCTGCTGCTGCTGCTGCTGCT 562
DB 515 AGAAGCCAGACTCAGGGAAGGAGGCTTCTTCTGAGTGGAGGCTTCCAGCTGGGC 574
QY 563 TTTTCAACATPCCAGGTGGCTATGAGATCCCTGAACTTTCTAAGTACCTGGATTTCACTCC 622
DB 575 AGACCTATGTGGATGCTGGATAGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 634
QY 623 ATGTCATGATATGACCTCCATGCTCTGGAGGCTTCCAGGAGGCTACCTGGGAGAGTAGTCTC 682
DB 635 ACCTTATGGCTACGACTTCCATGCTCTGGAGGAGGCTGCTGGAGGAGGCTGCTGGAGGAG 694
QY 683 TTTTCAAAATACCTACTGAGACTGGTAGCAATGCCCTACCTCAATCTGATGATGATGATGAT 742
DB 695 TCTTCAAGAGGCAAGAAGAGAGAGTGGTGCAGCAGCCAGCTTCAACGCTGGATGCTGCTGC 754
QY 743 ACTATTGGAGAACATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 802
DB 755 AACAGTGGCTGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 814
QY 803 GACACACCTTCACTCTGAGAAACCCCTCTGATAATGGAATGGTGGCCCTACCTCTGGTG 862
DB 815 GAGCTCTCTACACTGGCTCTCTATCAGACACAGAGTGGGGGCCCCAGCCACAGAGGT 874
QY 863 ATGGCCCTGCTGGGCCCTATACACAGAGGCTGGGTTCTGGGCCCTACTATGAGATTTGCA 922
DB 875 CTGGCACTCCAGGCCCTTTCAACCAAGGAAGAGGAGGATGCTGGCCCTACTATGAGTCTGCT 934
QY 923 CCTTCTGAGAAGTGGAGCCACTGAGTCTGGGATGCTGCCCTCCCAAGAGTGGCCCTATGCT 982
DB 935 CCT-----GGAAGGGGGGCCCAACAACAGAGATCCAGGATCAGAGAGTGGCCCTACATCT 988
QY 983 ATAAGGCCCAAGTGGCTTGGCTATGACAATATCAAGAGCTTCACTGTTTAAAGGCTCAGT 1042
DB 989 TCGGGACAAACAGTGGGTGGGCTTTGATGATGTTGGAGAGCTTCAAAACCAAGGTCAGCT 1048
QY 1043 GCTTAAAGCAGAACAAATTTGGAGGTGCCATGATCTGGGCCCTACTGACCTTGTGACTTCA 1102
DB 1049 ATCTGAAGCAGAGGAGGAGTGGGGGGCCATGCTCTGGGCACCTGGACTTAGATGACTTTG 1108
QY 1103 CTGGCTCTTCTGCTGATCAGGGAATAATTTCTCTGA 1138

Db 1109 CCGGCTTCTCTGCAACAGGCGCATACCCCTCA 1144

RESULT 5

US-09-151-011-3

; Sequence 3, Application US/09151011

; Patent No. 6057142

; GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: A Human Chitinase, Its Recombinant

; TITLE OF INVENTION: Production, Its Use For Decomposing Chitin, Its Use in

; TITLE OF INVENTION: Therapy or Prophylaxis Against Infection Diseases.

; NUMBER OF SEQUENCES: 16

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hoffmann & Baron, LLP

; STREET: 6900 Jericho Turnpike

; CITY: Syosset

; STATE: New York

; COUNTRY: United States of America

; ZIP: 11791

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/151,011

; FILING DATE: 10 - September - 1998

; ATTORNEY/AGENT INFORMATION:

; NAME: Morris, Robert C.

; REGISTRATION NUMBER: 42,910

; REFERENCE/DOCKET NUMBER: 294-32 DIV

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (516) 822-3550

; TELEFAX: (516) 822-3582

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1643 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: unknown

; TOPOLOGY: unknown

; MOLECULE TYPE: cDNA

; HYPOTHETICAL: NO

US-09-151-011-3

Query Match 28.0%; Score 426.8; DB 3; Length 1643;

Best Local Similarity 62.1%; Pred. No. 6.6e-123;

Matches 693; Conservative 0; Mismatches 417; Indels 6; Gaps 1;

QY 23 CAGGCTGTGCTCTTCTGCTGAATGCTCAGCTGGGCTCGCTACATCTGATGCTATT 82

Db 35 CAGGTTTCATGCTCTGCTGATGATCCATGGGCTCTCTCCAAACTGGTCTGCTACT 94

QY 83 TCACCAACTGGCCAGTATCGGCGAGGCTCTGGGAGCTTCAAGCCTGATGACATTAACC 142

Db 95 TCACCAACTGGCCAGTATCGGCGAGGCTCTGGGAGCTTCAAGCCTGATGACATTAACC 154

QY 143 CCGTCTGTGTTACTACCTGATCTATGCTTGGGATCGAGACAAATGATGACACCA 202

Db 155 CCAGCTTTGACCCACCTCATCTAGCCCTTCGCTGGCATGACCAACCACTGAGCA 214

QY 203 CCATGAATGGAATGATGTTACTCTCTATAAAGCTTTCAATGACTTGAAGGAGCA 262

Db 215 CCAGTGTGGAATGATGAGAGATCTCTACAGAGTTCAATGGCCTGAGAGATGAATC 274

QY 263 GCAACTGAAACCCCTCTGGCAATTTGGAGCTGGAACCTTTGGAATGCTCTCTTCACTA 322

Db 275 CCAAGCTGAAGACCCCTGTTAGCCATCGAGGCTGGAATTTGCGCACTCAGAGTTTCAAG 334

QY 323 CCATGCTTTCCACTTCTCAGAACCCGACACCTTCATTACCTCAGTCATCAATTTCTGC 382

Db 335 ATATGCTAGCCACGGCCAAACACCGTCAGACCTTTTGTCAACTCGGCCATCAGGTTTCTGC 394

QY 383 CTCAGTATGGGTTTGTGAGCTGGACCTGGACTGGGAATATACCCAGGCTCACCTGGGAGCC 442
Db 395 GCAAAATACAGCTTTTGACGGCTTGACCTTGAGTGGAGTACCCAGGAAGCCAGGGAGCC 454
QY 443 CTCCTCAGCAAGCATCTCTTCTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 502
Db 455 CTGCGGTAGACAGGAGGCGCTTCAAAACCTGTTACAGGACTTGGCCAAATGCTTCCAGC 514
QY 503 AGGAGGCTATTTCAGAGCAACAGGCCAGAGCTGATGTTACTGCTGCTGCTGCTGCTGCTG 562
Db 515 AGGAAGCCAGAGCTCAGGAAGGAGCGCTTCTTCTGAGTGCAGGGGTTCCAGCTGGGC 574
QY 563 TTTCCAAATCCAGGCTGGCTATGAGATCCCTGAACTTTCTTAAGTACCTGGAATTTATCC 622
Db 575 AGACCTATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 634
QY 623 ATGTCATGACATATGACCTCCATGCTCTCTGGAGGGCTACACTGGGAGAGATAGTCCTC 682
Db 635 ACCTTATGGCTTACGACTTCCATGCTCTTGGGAGAAAGGTCACGGGACATACAGCCGCC 694
QY 683 TTTACAAATPACCTTACTGAGACTGGTAGCAATGCTTACCTCAATGTGGATTTATGTCATGA 742
Db 695 TCTACAAGAGGCAAGAAGAGAGTGTGTCAGAGCCAGCTCAACGTGGATGCTGCTGTGTC 754
QY 743 ACTATTGGAAGAACCAATGGAGCCGCCAGCTGAGAAGCTCAATGTTGGATTCCAGAGATG 802
Db 755 AACAGTGGCTGCAGAGGGGACCCCTGCGCAGCAAGCTGATCTTGGCATGCTTACCTACG 814
QY 803 GACACACTTCACTCTGAGAAACCCCTCTGTAATTAAGAAATTTGGTGGCCCTTACCTCTG 862
Db 815 GACGCTCTCTACACTGGCTCTCTATCAGACACCAAGAGTGGGGGCCCGCCAGCAGGCT 874
QY 863 ATGGCCCTGCTGCGCTATACAGACAGGCTGGGTTCTGCGCTACTATGAGATTGCA 922
Db 875 CTGGCACTCCAGGCCCTTCCACAGGAAGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 934
QY 923 CCTTTCTGAGAACTGGAGCTGCTGAGGCTTGGGATGCTTCCCAAGAAAGTGCCTTATGCT 982
Db 935 CCT-----GGAAGGGGGCCACCAACAGAGAATCCAGGATCAGAAAGTGCCTTACATCT 988
QY 983 ATAAAGCCAAACAGAGTGGCTTGGCTATGACAATATCAAGAGCTTCACTGTTAAGGCTCAGT 1042
Db 989 TCCGGACAAACAGCTGGTGGCTTTGATGATGTTGGAGAGCTTCAAAACCAAGGTCAGCT 1048
QY 1043 GCTTAAGCAGAACAAATTTTGGAGTGCATGCTGGGCTGCTGCTGCTGCTGCTGCTGCT 1102
Db 1049 ATCTGAAGCAGAGGAGCTGGGGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1108
QY 1103 CTGGCTTTTCTGTATCAGGGAATAATTTCTCTCTGA 1138
Db 1109 CCGGCTTCTCTGCAACAGGCGCATACCCCTCA 1144

RESULT 6
US-09-343-623-3
; Sequence 3, Application US/09343623
; Patent No. 6303118
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: A human chitinase, its recombinant
; TITLE OF INVENTION: production, its use for decomposing chitin, its use
; TITLE OF INVENTION: in therapy or prophylaxis against infection diseases.
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11758
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
SOFTWARE:

CURRENT APPLICATION DATA: US/09/343, 623
FILING DATE:

PRIOR APPLICATION DATA: US/08/486, 839
FILING DATE: 07-June-1995

ATTORNEY/AGENT INFORMATION:

NAME: Baron, Ronald J.
REGISTRATION NUMBER: 29,281
REFERENCE/DOCKET NUMBER: 294-26

TELEPHONE: (516) 822-3550

TELEFAX: (516) 822-3582

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 1643 base pairs

TYPE: nucleic acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: CDNA

HYPOTHETICAL: NO

US-09-343-623-3

Query Match 28.0%; Score 426.8; DB 4; Length 1643;

Best Local Similarity 62.1%; Pred. No. 6.6e-123;

Mismatches 693; Conservative 0; Mismatches 417; Indels 6; Gaps 1;

23 CAGGTCTGGCTCTCTCTGCTGAATGCTCAGCTGGGGTCTGCTCAATCTGATGCTATT 82

35 CAGGTCTGATGCTCTGCTGATGATCCATGGGCTCTGCTCCAAACTGGTCTGCTACT 94

83 TCACCACTGGGCCCCAGTATCGCCAGGCTTGGGGAGCTTCAAGCTGATGATTAACC 142

95 TCACCACTGGGCCCCAGTATCGCCAGGCTTGGGGAGCTTCCCTCCCAAGGACTTGGACC 154

143 CTTGCTGTGCTACTCACTCTGATGCTATGCTCTGCTGGGATGAGCAATGAGATCACCA 202

155 CAGGCTTTGACCCACCTCATCTACGCTTCTGCTGATGAGCAATGAGATCACCA 214

203 CCATAGATGGAATGATGTTACTCTCTATAAGCTTTCAATGACTTTGAAAACAGGAACA 262

215 CCATGAGTGAATGAGGAGCTCTCTACAGGAGTTCAATGCTTCAATGAGGAGTGAATC 274

263 GCRAACTGAAACCTTCTGGCAATTTGGAGGCTTGGAACTTGGAACTGCTCTTCACTA 322

275 CCAAGCTGAAGACCTTGTAGCCATCGGAGGCTGGAATTCGGCACTCAGAAGTTCACAG 334

323 CCATGCTTTCCACTTCTCAGAACCGCCAGACCTTCACTCAGTCACTCAAAATTTCTGC 382

335 ATATGCTAGCCAGCCCAACACCTGACACCTTCTCACTCGGCCATCAGGTTTCTGC 394

383 GTCAGTATGGTTTGTATGAGCTGGACCTGGAGTGGGAATACCCAGGCTCACGTGGAGCC 442

395 GCAATACAGCTTTGACGGCTTGACCTTGACCTGGGAGTACCCAGGAAGCCAGGGAGCC 454

443 CTCCTCAGACAGCATCTCTCACHTCTGCTGTGTAAGGAATTCGCTGAGCTTTTGGAGC 502

455 CTGCCCTAGACAGGAGCGCTTTCACAACTCTGTACAGGACTTGGCCAAATGCTTCCAGC 514

503 AGGAGCTATTGAGACAACAGCCAGACCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 562

515 AGGAACCCAGACCTCAGGAGGAAGACGCTTCTCTGAGTGGCGGTTTCCAGCTGGGC 574

563 TTTCACATCCAGGCTGGCTATGATGCTTCACTTCTTCTAGTACCTGGATTTTCATCC 622

575 AGACCTATGTTGATGCTGATGAGGTTGAGCAAAATCCCAAGACCTGGATTTTGTCA 634

623 ATGTCATGATGATGATGCTTCCATGCTCTGGAGGCTACACTGGGGAGATGATCTTCT 682

635 ACCTTATGGCCTACGACTTCCACTGGCTTTGGGAGAGGTCACGGGAGATACAGCCCC 694

683 TTTACAAATACCTACTAGAGACTGGTAGCAATGCCTACCTCAATGTGATGATGATGATGA 742

695 TCTACAAAGGCAAGAGAGAGTGGTGACAGCCAGCCCTCAAGTGGATGCTGCTGCTGC 754

743 ACTATTGGAGAAATGGAGCCCGAGCTGAGAGCTCATTTGTTGGATTCCCAAGAGTATG 802

755 AACAGTGGCTGACAGAGGGGACCCCTGCCAGCAAGCTGATCTTTGGCATGCTACCTACG 814

803 GACACACTTCTATCTCTGAGAAACCCCTCTGATATGGAATGTTGGTCCCTACCTCTGGTG 862

815 GACGCTCTTTCACACTGGCTCTCTATCAGACACAGAGTGGGGGCCCCAGCCAGGGT 874

863 ATGGCCTGCTGGCGCTATACAGACAGGCTGGGTTCTGGGCTCTACTATGAGATTGCA 922

875 CTGGCACTCCAGCCCTTCCACCAAGGAGGAGTGTGGCTGCTACTATGAGTCTGCT 934

923 CTTTCTGAGAGTGGAGCCACTGAGGTCTGGGATGCTCTCCCAAGAGTCCCTATGCTCT 982

935 CCT-----GGAAGGGGGCCCAACAGAGATCCAGGATCAGAAGTCCCTACATCT 988

983 ATAAAGCCACAGTGGCTTGGCTATGACAAATATCAAGAGCTTCAAGTAAAGCTCAGT 1042

989 TCCGGGACACACAGTGGTGGGCTTTGATGATGAGAGCTTCAAAACCAAGGTCAGT 1048

1043 GGCTTAAGCAGAAATTTTGGAGTGGCCATGATCTGGGCAATGACCTTGTGATGACTTCA 1102

1049 ATCTGAGCAGAGGAGCTGGGGGGCCATGCTGGGCACTGGGACTTGAATGACTTTG 1108

1103 CTGGCTTTCTGATGATGAGGAAATTTCTCTCTGA 1138

1109 CCGGCTTCTCTCAACAGGGGCGGATACCCCTCA 1144

RESULT 7

US-08-486-839-5

Sequence 5, Application US/08486839

Patent No. 5928928

GENERAL INFORMATION:

APPLICANT:

TITLE OF INVENTION: A human chitinase, its recombinant

TITLE OF INVENTION: production, its use for decomposing chitin, its use

TITLE OF INVENTION: in therapy or prophylaxis against infection diseases.

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: Hoffmann & Baron

STREET: 350 Jericho Turnpike

CITY: Jericho

STATE: New York

COUNTRY: United States of America

ZIP: 11758

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/486,839

FILING DATE: 07 - June - 1995

ATTORNEY/AGENT INFORMATION:

NAME: Baron, Ronald J.

REGISTRATION NUMBER: 29,281

REFERENCE/DOCKET NUMBER: 294-26

TELEPHONE: (516) 822-3550

TELEFAX: (516) 822-3582

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 1713 base pairs

TYPE: nucleic acid

STRANDEDNESS: unknown

	Query Match	28.08; Score 426.8; DB 3; Length 1713;
	Best Local Similarity 62.1%; Pred. No. 6.8e-123;	
	Matches 693; Conservative 0; Mismatches 417; Indels 6; Gaps 1;	
QY	23 CAGGTCTGGCTCTTCTGCTGAATGCTCAGCTGGGGTCTGCTCAACTCTGATATGCTATT 82	
DB	+ + + + + + + +	
QY	35 CAGTTTCATGGTCTGCTGATGATCCATGGGGCTCTGCTCCAAACTGGTCTGCTACT 94	
DB	+ + + + + + + +	
QY	83 TCACCAACTGGGCCAGTATCGGCCAGGTCCTGGGAGCTTCAAGCCCTGATGACATTAACC 142	
DB	+ + + + + + + +	
QY	95 TCACCAACTGGGCCACTACAGACAGGGGAGGCTCGCTTCTGCCCAAGGACTTTGGACC 154	
DB	+ + + + + + + +	
QY	143 CTTGCGCTGTACTACCTGATCTATGCCCTTCTCTGGGATCGACAACTATGAGATCACCA 202	
DB	+ + + + + + + +	
QY	155 CCAGGCTTTGCACCCACTCTTCTACGCCCTTCGCTGGCATGACCAACCCACGCTGAGCA 214	
DB	+ + + + + + + +	
QY	203 CCATAGAAATGGAATGATGTACTCTCTATAAAGCTTTTCAATGACTTTGAAAAACAGGAACA 262	
DB	+ + + + + + + +	

Db 215 CCCTGAGTGAATGAGAGACTCTTACCAGGAGTTCAATGGCTGAAGAAGATGATC 274
Qy 263 GCAAACTGAAACCTCTCTGCAATTTGGAGCTGAACTTTGGAACTGCTCTTCACTA 322
Db 275 CCAAGCTGAGACCTGTAGCCATCGAGGCTGGAATTTGGGCACCTCAAGTTCACAG 334
Qy 323 CCATGGTTCCACTTCTCAGAACCGCCAGACCTTCTTACCTCAGTCAATCAAAATTTCTGC 382
Db 335 ATATGGTAGCCAGCCGCAACACCGTCAGACCTTTGTCAACTCGGCCATCAGTTTCTGC 394
Qy 383 GTGATGATGGTTGATGAGCTGACCTGAGCTGGGAATACCCAGGCTCAGTGGGAGCC 442
Db 395 GCAAAATACAGCTTTGACGGCTTACCTTGAAGTGGAGTACCCAGGAAGCCAGGGAGCC 454
Qy 443 CTCCTCAGGACAGCTCTCTTCACTGCTCTGTTGAAGAAATGCGTGAAGCTTTTCAGC 502
Db 455 CTGCGGTAGACAGGAGCGCTTCAACCTCTGTTACAGGACTTGGCCATGGCTTTCAGC 514
Qy 503 AGGAGGCTATTGAGACCAACAGGCCACAGCTGATGGTTACTGCTGTAGCTGAGTGGTGGGA 562
Db 515 AGGAAGCCAGACCTCAGGGAAGGAGCGCTTCTTCTGAGTCAGCGGTTCCAGCTGGC 574
Qy 563 TTTCACACATCCAGCTGGCTATGAGATCCCTGAATTTCTAAGTACCTGATTTTCATCC 622
Db 575 AGACCTATGTTGATGCTGATAGAGGTGGAGGTCGACAAATCGGCCAGAACCTGGATTGTCA 634
Qy 623 ATGTCATGACATGACCTCCATGCTCCATGCTGCTGAGTGGTACCTGAGTGGGAGAAATGCTC 682
Db 635 ACCTATGGCTCAGACTTCCATGCTCTTGGGAGGAGTGCAGGGACATTAAGAGCCCC 694
Qy 683 TTTACAAATACCTTCTGAGAGCTGATGAGTGAATGCTACCTCAATGTGATTTATGTCATGA 742
Db 695 TCTACAGAGGCAAGAGAGAGTGGTGCAGCAGCCAGCTCAAGTGGATGCTGCTGC 754
Qy 743 ACTATTGGAAGACATGAGGAGCCAGCTGAGAGCTCAATGTTGGATTTCCAGAGATG 802
Db 755 AACAGTGGCTGCAAGAGGAGGAGCCCTGCCAGCAAGCTGATCCTTGGCATGCCATACCTACG 814
Qy 803 GACACACCTTCATCTGAGAAACCCCTCTCATATGAAATGTTGGTGGCCCTACCTCTGGTG 862
Db 815 GACGCTCTTCACACTGGCTCTCATCAGACACAGAGTGGGGGCCACACAGGTT 874
Qy 863 ATGGCCCTGTGGCCCTATPACAGACAGCTGGGTTCTGGGCTTACTATGATGATTGCA 922
Db 875 CTGGCACTCCAGGCCCTTCAACCAAGGAAGGAGGATGCTGGCTTACTATGAGTCTGCT 934
Qy 923 CTTTCTGAGAGTGGAGCCACTGAGTCTGGATGCTTCCCAAGAGTCCCTATGCTTGCCT 982
Db 935 CCT-----GGAAGGGGGCCCAACAGAGATCCAGGATCAGAGGTTGCCCTACATCT 988
Qy 983 ATAAGGCCAACAGTGGCTTGGCTATGACAATATCAAGAGCTTTCAGTGTAAAGGCTCAGT 1042
Db 989 TCCGGACACACAGTGGTGGCTTGTATGATGTTGGAGAGCTTCAAAACCAAGCTCAGCT 1048
Qy 1043 GGCTTAAGCAGAACATTTTGGAGTGGCCATGATCTGGGCCATPAGCTTGTGATGACTTCA 1102
Db 1049 ATCTGAAGCAGAGGAGCTGGGGGGCCCATGCTCTGGGCACTGGACTTAGATGACTTTG 1108
Qy 1103 CTGGCTCTTCTGATCAGGGAATTTCTCTGA 1138
Db 1109 CCGGCTTCTCTGCAACAGGGGCCGATACCCCTCA 1144

RESULT 9

US-09-343-623-5

; Sequence 5, Application US/09343623

; Patent No. 6303118

; GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: A human chitinase, its recombinant

; TITLE OF INVENTION: production, its use for decomposing chitin, its use

; TITLE OF INVENTION: in therapy or prophylaxis against infection diseases.

NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann & Baron
STREET: 350 Jericho Turnpike
CITY: Jericho
STATE: New York
COUNTRY: United States of America
ZIP: 11758
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
SOFTWARE:
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/343,623
FILING DATE:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/08/486,839
FILING DATE: 07-June-1995
ATTORNEY/AGENT INFORMATION:
NAME: Baron, Ronald J.
REGISTRATION NUMBER: 29,281
REFERENCE/DOCKET NUMBER: 294-26
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 822-3550
TELEFAX: (516) 822-3582
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1713 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
US-09-343-623-5

Query Match 28.0%; Score 426.8; DB 4; Length 1713;

Best Local Similarity 62.1%; Pred. No. 6.8e-123;

Matches 693; Conservative 0; Mismatches 417; Indels 6; Gaps 1;

Qy 23 CAGTCTGGCTCTTCTGCTGAATGCTCAGCTGGGGTCTGCTCAATCTGATATGCTATT 82

Db 35 CAGTCTTCTGCTGCTGCTGATGCCATGGGGCTCTGCTCCAAAACCTGCTGCTACT 94

Qy 83 TCACCAACTGGGCCAGTATCGCCAGGCTTGGGGAGCTTCAAGCTGATGACATTAAC 142

Db 95 TCACCAACTGGGCCAGTATCGCCAGGCTTGGGGAGGCTGCTTCTGCTCCCAAGGACTTGAAC 154

Qy 143 CCTGCTGTGCTACTACCTGATCTATGCCCTTCTGCTGGATGCAAGAACATGATCAGCA 202

Db 155 CCAGCCTTTCACCCACCTCATCTAGCCTTCTGCTGCATGACCAACCACCTGATGCA 214

Qy 203 CCATAGAATGAATGATGTTACTCTCTATAAGCTTTCAATGACTTTGAAAAACAGGAACA 262

Db 215 CCATCAGTGAATGATGACGAGACTCTCTACAGGAGTTCAATGCTGCTGAAGAAGATGAATC 274

Qy 263 GCAAACTGAAACCTTCTGCGCAATTTGGAGCTGGAGCTTTGGAACCTGCTCTTCACTA 322

Db 275 CCAAGCTGAAGACCTTGTAGCCATCGGAGGCTGGAATTTCCGCACTCAGAAAGTTCCACAG 334

Qy 323 CCATGTTTCCACTTCTCAGAACCCGACACCTTCAATACCTCAGTCAATCAAAATTTCTGC 382

Db 335 ATATGGTAGCCAGGCCCAACACCGTCAGACCTTCTCACTGGGCCATCAGTTTCTGC 394

Qy 383 GTGATGATGGTTTGTATGAGCTGGAGCTGGAGTGGGAATACCCAGGCTCACCTGGGAGCC 442

Db 395 GCAAAATACAGCTTTGACGGCTTGAACCTTGAAGTGGAGTACCAGGAAGCCAGGGAGCC 454

Qy 443 CTCCTCAGCAGACATCTCTTCACTGCTCTGTTGAAGAAATGCTGAGCTTTTTCAGC 502

Db 455 CTGCGGTAGACAGAGGAGGCTTCAACACCTTGTACAGGACTTGGCCAATGCTTCCAGC 514

QY 503 AGGAGCTATTGAGACCAACAGGCCAGACTGATGTTTACGCTGCTGTAGCTGGTGGGA 562
Db 515 AGGAAGCCAGACCTCAGGAGGAAGAACCTTCTTCTGAGTGACGCGTTCCAGCTGGGC 574
QY 563 TTTCCAACTCCAGCTGGCTGATGAGATCCCTGAATCTTCTAAAGTACCTGATTTTCATCC 622
Db 575 AGACCTATGTGTGATGCTGGATACGAGGTGGACAAATCGCCAGAACCTGGATTTGTGCA 634
QY 623 ATGTCATGACATGATGACCTCCCTGCTGGAGGGCTACACTGGGAGAAATAGTCTCTC 682
Db 635 ACCATTATGCTACGACTTCCATGGCTTCTGGAGAGGTACGGGACATACAGCCCCC 694
QY 683 TTTTAAATACCTACTGAGACTGGTAGCAATGCTTCAATGATGATGATGATGATGATG 742
Db 695 TCTAAGAGCAAGAGAGAGTGTGAGAGGTGGAGGTGGAGGTGGAGGTGGAGGTGGAG 754
QY 743 ACTATTGAAGCAATGAGAGCCAGCTGAGAGCTGATGTTGGATGCCAGAGATG 802
Db 755 AACAGTGTGCTGAGAGGGGACCTGCTGCAAGAGCTGATCTTGGCATGCTTACCTACG 814
QY 803 GACACACCTTCTATGAGAACCCCTCTGATAATGGAATGGTGGCCCTTACCTCTGGTG 862
Db 815 GACGCTCTTCTACACTGCTCTCTATCAGACACAGAGTGGGGCCCGCCAGCAGGT 874
QY 863 ATGGCCCTGCTGGCCCTATACAGAGAGGTGGGTCTGGGCCCTTACTATGAGATTTGCA 922
Db 875 CTGGCACTCCAGGCCCTTCAACAGAGAGAGGAGGTGCTGGCTTACTATGAAGTCTGCT 934
QY 923 CTTTCTGAGAGTGGAGCCACTGAGTCTGGGATGCTCCCAAGAGTGGCCCTATGCT 982
Db 935 CCT-----GGAAGGGGGCCCAACAGAGATCCAGGATCAGAGGTGGCCCTACATCT 988
QY 983 ATAAGGCCCAAGTGGCTTGGCTATGACATATCAAGAGCTTTCAGTGTAAAGGCTCAGT 1042
Db 989 TCCGGGCAACAGTGGTGGCTTGGCTTGGATGATGAGAGCTTCAAAACCAAGGTGAGT 1048
QY 1043 GGCTTAAGCAGAACAAATTTGGAGTGGCCATGATGCTGGCCATGACCTTGTATGACTTCA 1102
Db 1049 ATCTGAAGCAAGAGGAGTGGGGGGGCTGCTGGGCACTGGACTTATGATGACTTTG 1108
QY 1103 CTGGCTTTCTGATCAGGAGAAATTTCTCTGCA 1138
Db 1109 CGGCTTCTCTGCAACAGGGCGGATACCCCTCA 1144

RESULT 10
US-09-039-198A-3
; Sequence 3, Application US/09039198A
; Patent No. 6200951
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; APPLICANT: Tjoelker, Larry W.
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive/6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/039.198A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rln-Laures, Li-Hsien
; REGISTRATION NUMBER: 33.547

REFERENCE/DOCKET NUMBER: 27866/34391
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1656 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 27..1424
; NAME/KEY: mat_peptide
; LOCATION: 90..1424
US-09-039-198A-3
Query Match 27.9%; Score 425.2; DB 4; Length 1656;
Best Local Similarity 62.0%; Pred. No. 2.1e-122;
Matches 692; Conservative 0; Mismatches 418; Indels 6; Gaps 1;
QY 23 CAGGTCTGGCTCTTCTGCTGATGCTCAGCTGGGTCTGCTGCTACAAATCTGATGATGCTATT 82
Db 49 CAGGTCTGATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 108
QY 83 TCACCAACTGGCCAGTATCGGCAGGCTCTGGGAGGCTTCAAGCTGATGATGATGATGATGATGAT 142
Db 109 TCACCAACTGGCCAGTATCGGCAGGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 168
QY 143 CTG 202
Db 169 CCAGCTTTGACCCACCTCATCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 228
QY 203 CCATAGAT 262
Db 229 CCACCTGAT 288
QY 263 GCAAACTGAAACCTCTCTGCAATTTGGAGGCTGGAACCTTTGGAACCTGCTCTCTTTCACAT 322
Db 289 CCAAGCTGAAGACCTCTGTTAGGCTGAGGCTGGAATTTGAGCTGAGAGCTGAGAGCTGAGAG 348
QY 323 CCATGCTTTCCACTTCTCAGAACCCAGACCTTCTTACCTCAGTCAATCAATTTCTGCT 382
Db 349 ATATGCTAGCCACGCGCAACACCTGTCAGACCTTTGTCACTCGGCTCAGCTGCTGCTGCTGCT 408
QY 383 GTCAGTATGGGTTTATGAGCTGAGCTGGAGCTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 442
Db 409 GCAATATACAGCTTTGACGCGCTTGACCTTGAAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 468
QY 443 CTCTCTCAGCAACAGCATCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 502
Db 469 CTGCGCTAGACAAAGGAGGCTTCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 528
QY 503 AGGAGGCTATTGAGAGCAACAGGCCAGACTGATGATGATGATGATGATGATGATGATGATGATGAT 562
Db 529 AGGAAGCCAGACCTCAGGGAAGGAGGCTTCTTCTGAGTGCAGCGGTTCCAGCTGGGCG 588
QY 563 TTTTCAACATCCAGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 622
Db 589 AGACCTATGCTGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 648
QY 623 ATGTCATGACATATGACCTTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 682
Db 649 ACCTTATGCTTACGACTTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 708
QY 683 TTTTCAAAATACCTTACTGAGACTGGTAGCAATGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 742
Db 709 TCTAAGAGGCAAGAGAGTGGTGGCAGCAGCAGCTCAACCTGATGATGATGATGATGATGATGAT 768
QY 743 ACTATTGGAAGCAATGAGGCCAGCTGAGAGCTCAATGCTGATGATGATGATGATGATGATGATGAT 802

Db 769 AACAGTGGCTGCAGAGGGGACCCCTGCCAGCAAGCTGATCTTTGGCATGCCCTACCTACG 828
QY 803 GACACACCTTCATCTCTGAGAAACCCCTCTGATATGAAATGCTGCTTACCTACCTCTGGTG 862
Db 829 GAGCTCTCTACACCTGGCTCTCTATCAGACACAGAGTGGGGCCCGCCAGCAGAGGT 888
QY 863 ATGGCCCTGCTGGCGCTATACACAGACAGCTGGGTCTGGGCTTACTATGAGATTGCA 922
Db 889 CTGGCACTCCAGGCCCTTCCACCAAGAGGAGGATGCTGGCTACTATGAGTCTGCT 948
QY 923 CTTTCTGAGAACTGAGGACCTGAGGTCTGGGATGCTCCCAAGAGTGGCCCTATGCT 982
Db 949 CTT-----GGAAGGGGGCCACCAACAGAGATCCAGGATCAGAGGTGGCCCTACATCT 1002
QY 983 ATAAAGCCACAGCTGGCTGCTATGACAAATATCAAGAGCTTCAGTGTAAAGCTCAGT 1042
Db 1003 TCCGGACACACCTAGTGGGTGGCTTGTATGATGAGAGCTTCAAAACCAAGGTGAGCT 1062
QY 1043 GGCTTAAGCAGAAATTTTGGAGTGCCATGATCTGGGCCATTTGACCTTGTATGACTTCA 1102
Db 1063 ATCTGAAGCAGAGGACTGGCGGGCCATGCTGGGCACTGGACTTAGATGACTTGG 1122
QY 1103 CTGGCTCTTCTGATCAGGGAATTTCTCTCTGA 1138
Db 1123 CCGGCTTCTCTGCAACAGGCGCGATACCCCTCA 1158

RESULT 11
US-877-599-3
; Sequence 3, Application US/0877599
; Patent No. 6372212
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; TITLE OF INVENTION: Chitinase Materials and Methods
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE:
; PRIORITY APPLICATION DATA:
; CLASSIFICATION: 514
; APPLICATION NUMBER: US 08/663,618
; FILING DATE: 14-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 27866/33994
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1656 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 27..1424
; FEATURE:

; NAME/KEY: mat_peptide
; LOCATION: 90..1424
US-08-877-599-3
Query Match 27.9%; Score 425.2; DB 4; Length 1656;
Best Local Similarity 62.0%; Pred. No. 2.1e-122;
Matches 692; Conservative 0; Mismatches 418; Indels 6; Gaps 1;
QY 23 CAGGTCTGGCTCTTCTGCTGAATGCTCAGCTGGGTCTGCTCACAATCTCATATGCTATT 82
Db 49 CAGGTCTGATGGTCTGCTGATGATCCCATGGGCTCTGCTCAAAATCTGCTGCTACT 108
QY 83 TCACCAACTGGGCCAGGTATCGGCCAGGTCTGGGAGCTTCAAGCTCTGATGACATTAAC 142
Db 109 TCACCAACTGGGCCAGGTATCAGACAGGGGAGGCTCGCTTCTGCCCCAAGACTTGAAC 168
QY 143 CTGCTCTGATCTACCTGATCTATGCTCTGCTGGGATGACAGAAATGATGATCAACCA 202
Db 169 CCAGCTTTTGCACCCACCTCATCTAGCTCTGCTGCTGATGACCAACACCAAGCTGAGCA 228
QY 203 CCATAGAAATGATGATGTTACTCTATATAAGCTTTCAATGACTTGAATAACAGGAACA 262
Db 229 CCACTGAGTGAATGACGAGACTCTCTACAGGAGTTCATAGCTCTCAAGAGATGATC 288
QY 263 GCAAACTGAAACCCCTCTGGCAATTTGGAGCTTGAACCTTTGGAACCTGCTCTTCACTA 322
Db 289 CCAAGCTGAAGACCTTGTAGCCATCGGAGGCTGGAATTTTCAGCACTCAGAAGTTCACAG 348
QY 323 CCATGTTTCCACTTCTCAGAACCCGACAGACTTCACTACCTCAGTCAATCAATTTTCTGC 382
Db 349 ATATGTTAGCCACGGCCAAACACCTGTCAGACTTGTCAACTCGGCATCAGGTTTCTGC 408
QY 383 GTCAGTATGGTGTGATGGACTGGCTGGAGTGGGAATACCAGGCTCAGCTGGGAGCC 442
Db 409 GCAATACAGCTTTGACGGCTTGACCTGACTGGGAGTACCAGGAGCCAGGAGCC 468
QY 443 CTCTCAGGACAAAGCATCTCTTCACTGCTCTGCTGAGGAAATGCGTGAAGCTTTTGAAGC 502
Db 469 CTGCGCTAGACAAAGGAGCGCTTCACAACTCTGTCAGAGACTTTGGCCAAATGCTTCAGC 528
QY 503 AGGAGCTATTGAGAGCAACAGCCCGACAGCTGATGTTACTGCTGCTGATGCTGGGGA 562
Db 529 AGGAAGCCAGACTCAGGGAAGAGAGCTTCTCTGAGTGCAGCGGTTCCAGCTGGGC 588
QY 563 TTTTCAACATCCAGCTGCTGATGATGATCCCTGAGATCTTCTAAGTACCTGATTTTCTGC 622
Db 589 AGACCTATGTTGATGCTGGATACAGAGTGGACAAATTCGCCAGAACCTGATTTTCTCA 648
QY 623 ATGTCATGATATGACCTCCATGCTCTGCTGGAGGCTTACACTGGGAGAAATAGTCTCTC 682
Db 649 ACCTTATGGCTACGACTTCCATGGCTCTTGGGAGAGGTCAGCGGACATAACAGCCCC 708
QY 683 TTTTCAAAATACCTACTGAGACTGGTAGCAATCCCTACCTCAATGCTGATGATGATGCA 742
Db 709 TCTACAAGAGGCAAGAGAGAGTGGTGCAGCAGCCCTCAACGTTGATGCTGCTGTC 768
QY 743 ACTATTGGAAGAAATGAGCCCGAGCTGAGAGCTGAGAGCTTCTGATTTGATTTCCAGAGATG 802
Db 769 ACAGTGGCTGACAGAGGGGACCCCTGCCAGCAAGCTGATCTCTTGGCATGCCCTACCTACG 828
QY 803 GACACACTTTCATCTCAGAAACCCCTCTGATGATGATGATGATGATGATGATGATGATGATG 862
Db 829 GAGCTCTCTACACTGGGCTCTCTCATCAGACACAGAGTGGGGGGGGGGGGGGGGGGGGGG 888
QY 863 ATGGCCCTGCTGGCGCTTATACAGACAGCTGGGTCTGGGCTTCTGGGCTTACTATGATGATTTGCA 922
Db 889 CTGGCACTCCAGGCCCTTCCACCAAGAGGAGGAGTGGTGGCTTCTGCTTCTGATGATGATGAT 948
QY 923 CTTTCTGAGAGTGGAGCCACTGAGGTCTGGGATGCTGCCCAAGAGTGGCTTCTGATGATGATG 982
Db 949 CCT-----GGAAGGGGGCCACCAACAGAGATCCAGGATCAGAGGTGGCCCTACATCT 1002
QY 983 ATAAAGCCACAGTGGCTTGGCTATGACAAATATCAAGAGCTTCAGTGTAAAGCTCAGT 1042

Db	469	CTGCGCTAGACAAAGGAGCGCTTCACAACCTGGTCAGAGACTTGGCCCAATGCCCTCCAGC	528
QY	503	AGGAGGCTATTGAGAGCAACAGGCCAGACTGATGTTACTGCTGCTAGCTGTGGGA	562
Db	529	AGGAAGCCAGACCTCAGGGAAGGACGCTTCTTCTGAGTCAGCGGTTCCAGCTGGGC	588
QY	563	TTTTCACAATCCAGGTGGCTATGAGATCCCTGAACCTTCTTAAGTACCTGGATTTTCATCC	622
Db	589	AGACCTTATGTGATGCTGATACGAGGTGCACAAAATCGCCAGAACCTGGATTTTGTGCA	648

649 ACCCTATGGGCTACGACCTTCATGGCTCTGGGAGAAGTCAAGGACATAACAGCCCCC 708

683	QY	TTTTCAAATACCCCTACTGAGACTGGTAGCAATGCCCTACCTCAATGTGGATTATGTGCATGA	742
709	Db	TCTACAGAAGGCAAGAAGAGAGTGGTGCAGACGCCAGCCCTCAACGTGGGATGCTGCTGTGC	768
743	QY	ACTATTGGAGAACAATGAGCCCCCAGCTGAGAAAGCTCATTTGGATTCCCCAGAGATAG	802
769	Db	AACAGTGGCTGCAAGAGGGGACCCCTGCCAGCAAGCTGATCCTTGGCATGGCTTACCTACG	828
803	QY	GACACACCTTCATCCTTGAAACCCCTCTGATAATGGAATTGGTGGCCCCCTACCTCTGGTG	862
829	Db	GAGCTCCTTCACACTGGCCTCCTCATCAGACACCAGAGTGGGGGCCCCAGGCCACAGGGT	888
863	QY	ATGGCCCTCTGGCGCCTATACCAGACAGCTGGGTTCCTGGGCCTACTATGAGATTGCA	922
889	Db	CTGGCACTCCAGCCCTTCACCCAGGAGGAGGATGCTGGCCTACTATGAGTCTGCT	948
923	QY	CTTTTGTGAAGTGGAGCCACTGAGTCTGGGATGCCTCCCAAGAAGTGCCCTATGCCT	982
949	Db	CCT-----GGAGGGGGCCACCAACAGAGATCCAGGATCAGAAGGTGCCCTACATCT	1002
983	QY	ATAAGGCCAACGAGTGGCTGGCTATGACAATATCAAGAGCTTCAGTGTTAAGGCTCAGT	1042
1003	Db	TCCGGGACACCAAGTGGGTGGGCTTTTGATGATGTGGAGAGCTTCAAAACCAAGGTCAGCT	1062

```

04Y      1063 ATCTGAAGCAGAAGGACTGGCGGGGCCCATGTCTGGGCACCTGGACTTAGATGACTTTG 1122
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
04Y      1103 CTGGCTCTTCTCTGATCAGGGAATAATTTCCTCTGA 1138
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
04Y      1123 CCGGCTCTTCTGCAACCAGGGCCGATACCCCCTCA 1158
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 13
US-08-694-915-1
; Sequence 1, Application US/08694915
; Patent No. 5811535
; GENERAL INFORMATION:
; APPLICANT: Adamou, Julie
; APPLICANT: Kirkpatrick, Robert
; APPLICANT: Rosenberg, Martin
; TITLE OF INVENTION: HUMAN CARTILAGE GP39-LIKE GENE
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-2799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/694,915
; FILING DATE:
; CLASSIFICATION:
```


INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1526 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
US-08-694-915-3

Query Match 21.7%; Score 330.8; DB 1; Length 1526;
Best Local Similarity 58.9%; Pred. No. 5.5e-93;
Matches 657; Conservative 0; Mismatches 432; Indels 27; Gaps 4;

QY 55 GGGCTGCTCAACATCGATATGCTATTTACCAACTGGGCCAGTATCGGCCAGGTCTG 114
DB 199 GGATCTGGCTTCAAACTGGTTTGTACTTTTACCAACTGGTCCAGGACCGGCAACCA 258

QY 115 GGGAGCTCAAGCCTGATGACATTAACCCCTGCTGTCTACTCACTCATCTATGCTCTT 174
DB 259 GGAATTCACCCCTGAGAATATTGACCCCTTCTCTATCTCTCATCTCTATTCATTC 318

QY 175 GCTGGGATCGAGAACAATGAGATCACCACTAGATGAATGGAATGATGTTTACTCTCTATAA 234
DB 319 GCCAGCATCGAACAACAAGGTATCATCAAGGACAGAGTGAAGTATGCTCTACACAG 378

QY 235 GCTTTCATGACTTGAAGAACAGGAACAGCAAGCAACTGAAACCCCTCGCTGGCAATGGAGGC 294
DB 379 ACCATCAACAGCTCAAAACCAAGAATCCCAAACTGAAATTTCTTCTCCATTTGGAGGG 438

QY 295 TGGAACTTTGAACTGCTCTCTTCTACCTACCATGTTTCCACTTCTCAAGACCGCAGACC 354
DB 439 TACCTGTTTGTTCAAAGGGTTTCCACCTATGTTGGATTTCTTCAATACAGCTGTGAA 498

QY 355 TTCATTACCTCAGTCATCAAAATTTCTGCTGATGTTGATGGTCTGACCTGGAC 414
DB 499 TTCATTAACTCCATAATCCTTCTGAGGAACCAATACTTTGATGGACTGATGATGAGC 558

QY 415 TGGNAATACCAAGCTCAGTGGAGCCCTCTCAGGACAGATCTCTTCACTGCTCTG 474
DB 559 TGGATCTACCCAGA-----TCAGAAAGAAACACTCATTTCACTGCTGCTG 603

QY 475 GTGAAGAAATGCGTGAAGCTTTTGAAGCAGAGGCTATTGAGACCAAGCCAGACTG 534
DB 604 ATTCATGAGTTAGCAGAGCCCTTTCAGAGGACTTCACAAATCCCAAGAAAGGCTT 663

QY 535 ATGTTACTGCTGTAGCTGAGTGGGATTTTCAACATCCAGGCTGGCTATGAGATCCCT 594
DB 664 CTCTTGACTCGGGGTATCTGAGGAGGCAATGATTGATAACAGCTATCAAGTTGAG 723

QY 595 GAACCTTCTAGTACCTGGATTTTCATCATGTCATGACATATGACCTCCATGGCTCTGG 654
DB 724 AAACCTGCAAAAGATCTGGAATTTTCATCAACCTCTGCTCTTCTGACTTCATGGTCTGG 783

QY 655 GA-----GGGCTACACTGGGAGATAGTCTCTTTTACAAATACCTTACTGAGACTGT 708
DB 784 GAAAGCCCTTATCACTGGCCACACAGCCCTCTGAGCAAGGGTGGCAGGACAGAGGG 843

QY 709 AGCAATGCTACCTCAATGTGGATTTATGTCATGAATTTTGAAGAAACAATGGAGCCCA 768
DB 844 CCAAGCTCTACTACAATGTGGAATATGCTGTGGGTACTGATACATAGGAATGCA 903

QY 769 CTGAGAGCTCATTGTTGATTCAGGATATGAGACACACTTCATCTTCAAGAAACCC 828
DB 904 TCAGAGAGGTGGTCATGGGATCCCAATATGGGCACTCTCTCACACTG---GCCCT 960

QY 829 TCTGATATGAAATTTGGTCCCTTACCTCTGATGCTGCTGCTGGCCCTATACACAGA 888
DB 961 GCAGAAACCCCTGGGGGCCCTTGCCTCTGGCCCTGGAGCTGTGAGCCCATCACAG 1020

QY 889 CAGGCTGGTCTTGGGCCCTACTATGAGATTTGCACCTTCTTGAGAGTGGAGCCACTGAG 948
DB 1021 TCTTCAGGCTTCCCTGGCCTATTATGAGATTCGCCAGTTTCTTCTGAAA---GGAGCCCAAGATC 1077

QY 949 GTCTGGGATGCTCCCAAGAGTGGCCCTATGCTATAGGCAAGCAGTGGCTTGGCTAT 1008
DB 1078 ACGGGCTCCAGGATCAGCAGGTTCCCTACGAGTCAAGGGGAAACAGCTGGGTGGCTAT 1137

QY 1009 GACAATATCAAGAGCTTCAGTGTAAAGCTCAGTGGCTTAAGCAAGCAAAATTTTGGAGGT 1068
DB 1138 GATGATGTAAGAGTATGGAGCAAGGTTCAAGTTCTTAAGAAATTTAAACCTGGAGGA 1197

QY 1069 GCCATGATCTGGGCAATTCACCTTGATGACTTCACTGGCTCTTCTGTCATCAGGAAAA 1128
DB 1198 GCCATGATCTGCTTATTTGACATGATGACTTCACTGGCAAACTCTGCAACCAAGGCCCT 1257

QY 1129 TTTCTCTGACTTCTACTTTTGAACAAAGCCCTTGGC 1164
DB 1258 TACCTCTTGTCCAAAGCAGTCAAGAGAACGCTTGGC 1293

RESULT 15
PCT-US94-07754-4
; Sequence 4, Application PC/TUS9407754
; GENERAL INFORMATION:
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
; TITLE OF INVENTION: ASSAY FOR YKL-40 AS A MARKER FOR
; TITLE OF INVENTION: DEGRADATION OF MAMMALIAN CONNECTIVE TISSUE MATRICES
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SPENSLEY HORN JUBAS & LUBITZ
; STREET: 1800 CENTURY PARK EAST, FIFTH FLOOR
; CITY: LOS ANGELES
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/07754
; FILING DATE: 08-JUL-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: HOWELLS, STACY L.
; REGISTRATION NUMBER: 34,842
; REFERENCE/DOCKET NUMBER: FD 3665
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/455-5100
; TELEFAX: 619/455-5110
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1681 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: YKL-40
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 135..1681
PCT-US94-07754-4

Query Match 20.6%; Score 314.8; DB 5; Length 1681;
Best Local Similarity 58.2%; Pred. No. 5.7e-88;
Matches 641; Conservative 0; Mismatches 437; Indels 24; Gaps 4;

QY 22 ACAGTCTGGCTCTTCTGCTGCTGAATGCTCAGTGGGTGCTGCTACAACTCTGATGCTAT 81
DB 93 ACAGCTTTGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTAC 152

Search completed: July 3, 2003, 08:43:20
Job time : 75.5873 secs

```
QY 82 TTCACCAACTGGGCCAGTATCGCCAGGTCTCTGGGAGCTTCAAGCCTGATGACATTAC 141
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 153 TACACAGCTGGTCCAGTACCGGAGGAGGAGTGGAGCTGCTCCAGATGCCCTTGAC 212
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 142 CCGTGGCTGTGACTACCTGATCTATGCCCTTTGCTGGGATGAGAACAAATGATCACC 201
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 213 CGCTTCTGTGTACCCACATCATCTACAGCTTTGCCAATATAAGCAACGATCACATCGAC 272
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 202 ACCATAGATGGAATGATGTTACTCTCTATAAAGCTTTCAATGACTTGRAAAACAGGAAC 261
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 273 ACCTGGAGTGAATGATGAGCGCTCTACGGCATGCTCAACACACTCAACAAACAGAAC 332
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 262 AGCAAACTGAAACCTCTCTGGCAATTTGGAGGCTGGAACCTTTGGAACCTGCTCTTCACT 321
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 333 CCCAACTGAAGACTCTCTTGTCTGCGGAGGATGGAACCTTTGGGTCTCAAGATTTTCC 392
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 322 ACCATGGTTTCCACTTCTCAGAACCGCCAGACCTTCAATACCTCAGTCATCAATTTCTG 381
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 393 AAGATAGCTTCCAAACCCAGAGTCGCGGACTTTTCATCAAGTCAGTACCGGCATTTCTG 452
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 382 CGTCAGTATGGGTTTGTGAGCTGGACCTGGACTGGGAATACCCAGGCTCACGTGGGAGC 441
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 453 CGCACCCATGGCTTTGATGGCGTGACCTTGCTGCTGCTTACCTGGACGGAGA----- 506
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 442 CTTCTCAGACAAAGCATCTTTCACTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 501
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 507 -----GACAAACACCATTTTACCACCCCTAATCAAGGAATGAAGGCCGAATTTATA 557
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 502 CAGGAGGCTATTGAGAGCAACAGGCCAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 561
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 558 AGGAAGCCAGCCAGCGGAAAAAGC---AGCTCTGCTCAGCGCAGCAGCTGTCTGCGGGG 614
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 562 ATTTCCAACTCAGGCTGGCTGATGATCCCTGAACTTTTAACTACCTGGATTTTCATC 621
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 615 AAGTCACTATGACAGCAGCTATGATGCCAAGATATCCCAACACCTGGATTTCAAT 674
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 622 CATGTCATGACATATACCTTCCATGCTCTGGGAGGCTTACACTGGGGAGAAATAGTCTCT 681
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 675 AGCATCATGACCTTACCATTTTCATGGCGCTGGCGTGGGACACAGGCCATCACAGTCCC 734
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 682 CTTTACAATACCTACTGAGACTGTAGCAATGCTTACCTCAATGTGGATATATGTCATG 741
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 735 CTGTTCCGAGGTGAGGAGTCAAGTCTCTGACAGATTCAAGCAACACTGACTATGCTGTG 794
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 742 AACTATTGGAAGAACAAATGGAGCCCGCAGCTGAGAAGCTCATTTGTTGATTCCAGAGAT 801
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 795 GGTATCATGTTGAGGCTGGGGCTCTGCCAGTAACTGGTGGTGGCATCCCCACCTTC 854
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 802 GGACACACCTTCACTCTGAGAAACCCCTCTGATAATGGAATGGTGGCCCTACCTCTGCT 861
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 855 GGGAGGAGCTTCACTCTG---GCTTCTTCTGAGACTGGTGTCCAGCGCCAATCTCAGGA 911
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 862 GATGGCCCTGCTGGCCCTTATACCAGACAGGCTGGGTCTTGGGCCCTACTATGAGATTTGC 921
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 912 CCGGGAAATTCAGGCGCGTTACCAAGGAGGAGGAGCCCTTGCCTACTATGAGATCTGT 971
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 922 ACCTTTCTGAGAAGTGGAGCACTGAGGTCTGGGATGCTCCCAAGAGTGCCCTATGCG 981
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 972 GACTTCT---CCGGAGCCACAGTCCATAGAACCTCGGCAGCAGGTCCCTATGCC 1028
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 982 TATAAGGCCAAGAGTGGCTTGGCTATGACAATATCAAGAGCTTCACTGTTAAGGCTCAG 1041
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1029 ACCAAGGGCAACAGTGGTAGGATACGACGACCAAGGAAAGGCTCAAAAGCAAGGTGCAG 1088
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1042 TGGCTTAACGACAGACAAATTTGGAGTGCCATGATCTGGGCCATTTGACCTTGTGACTTC 1101
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1089 TACCTGAAGGATAGGAGCTGGCAGGCCCATGGTATGGGCCCTGGACCTGGATGACTTC 1148
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1102 ACTGGCTCTTCTGTGATCAGG 1123
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1149 CAGGGCTCTTCTGCGGCCAGG 1170
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	1525	100.0	1525	9	US-10-004-2198-3	Sequence 3, Appli
2	1041	68.3	1625	9	US-10-004-2198-2	Sequence 2, Appli
3	430	28.2	1038	9	US-10-202-438A-14	Sequence 14, Appli
4	327.6	21.5	1391	10	US-09-822-830A-402	Sequence 402, Appli
5	327.2	21.5	449	10	US-09-960-352-678	Sequence 678, Appli
6	319.6	21.0	1925	9	US-10-097-340-44	Sequence 44, Appli
7	311.6	20.4	1681	10	US-09-215-077A-4	Sequence 4, Appli
8	311.6	20.4	1681	10	US-09-262-213A-4	Sequence 4, Appli
9	259.8	17.0	1474	10	US-09-765-231A-8	Sequence 8, Appli
10	152.2	10.0	415	10	US-09-960-352-2589	Sequence 2589, A
11	149.6	9.8	414	10	US-09-960-352-13050	Sequence 13050, A
12	148	9.7	410	10	US-09-960-352-7364	Sequence 7364, A
13	148	9.7	424	10	US-09-960-352-3057	Sequence 3057, A
14	147.4	9.7	418	10	US-09-960-352-3072	Sequence 3072, A
c 15	136.2	8.9	399	10	US-09-960-352-265	Sequence 265, Appli
c 16	134.8	8.8	411	10	US-09-960-352-14919	Sequence 14919, A
c 17	127.8	8.4	384	10	US-09-960-352-5721	Sequence 5721, A
c 18	124.6	8.2	398	10	US-09-960-352-14460	Sequence 14460, A
c 19	121.4	8.0	1608	9	US-10-218-743-20	Sequence 20, Appli

```
Db 61 GCCTACATCTGATATGCTATTTTCCAACTGGGCCAGTATCGGCCAGGTCTGGGAGC 120
QY 121 TTCAAGCTGATGACATTAACCCCTGCTGTGTACTACCTGATCTATGCTTTGCTGG 180
Db 121 TTCAAGCTGATGACATTAACCCCTGCTGTGTACTACCTGATCTATGCTTTGCTGG 180
QY 181 ATCAGAACATGAGATCACCACCATAGAAATGGAATGATGTTACTCTCTATAAAGCTTC 240
Db 181 ATCAGAACATGAGATCACCACCATAGAAATGGAATGATGTTACTCTCTATAAAGCTTC 240
QY 241 AATGACTTGAAGAACAGGACAGCAAACTGAAACCCCTCTGCAANTGGAGGCTGGAAC 300
Db 241 AATGACTTGAAGAACAGGACAGCAAACTGAAACCCCTCTGCAANTGGAGGCTGGAAC 300
QY 301 TTGGAACCTGCTCTTCACTACCATGTTTCCACTTCTCAGAACGCCGACAGCTTCATT 360
Db 301 TTGGAACCTGCTCTTCACTACCATGTTTCCACTTCTCAGAACGCCGACAGCTTCATT 360
QY 361 ACCTCAGTCATCAAAATTTCTGCTCAGTATGGGTTTGTAGGACTGGAGCTGGGA 420
Db 361 ACCTCAGTCATCAAAATTTCTGCTCAGTATGGGTTTGTAGGACTGGAGCTGGGA 420
QY 421 TACCCAGGCTCAGTGGGAGCCCTCTCAGGACAAGCATCTTCACTGTCTGTGTGAAG 480
Db 421 TACCCAGGCTCAGTGGGAGCCCTCTCAGGACAAGCATCTTCACTGTCTGTGTGAAG 480
QY 481 GAAATCGGTGAAGCTTTTGGAGAGGAGCTATTGAGAGCAACAGGCCAGACTGATGGTT 540
Db 481 GAAATCGGTGAAGCTTTTGGAGAGGAGCTATTGAGAGCAACAGGCCAGACTGATGGTT 540
QY 541 ACTGCTGTGATGCTGGGATTTCCACATCCAGGCTGGCTATGAGATCCCTGAACTT 600
Db 541 ACTGCTGTGATGCTGGGATTTCCACATCCAGGCTGGCTATGAGATCCCTGAACTT 600
QY 601 TCTAAGTACTGATTTTCACTCATGTCATGACATATGACTCCCTGCTGGAGGCT 660
Db 601 TCTAAGTACTGATTTTCACTCATGTCATGACATATGACTCCCTGCTGGAGGCT 660
QY 661 TACACTGGGAGAAATAGTCTCTTTTACAATACCCTACTGAGACTGGTAGCAATGCTAC 720
Db 661 TACACTGGGAGAAATAGTCTCTTTTACAATACCCTACTGAGACTGGTAGCAATGCTAC 720
QY 721 CTCATGTGGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
Db 721 CTCATGTGGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
QY 781 ATTGTTGGATTTCCAGAGTATGACACACTTCTCATCTGAGAACCCCTCTGATATGGA 840
Db 781 ATTGTTGGATTTCCAGAGTATGACACACTTCTCATCTGAGAACCCCTCTGATATGGA 840
QY 841 ATGTTGCCCCCTACCTCTGATGATGATGATGATGATGATGATGATGATGATGATG 900
Db 841 ATGTTGCCCCCTACCTCTGATGATGATGATGATGATGATGATGATGATGATGATG 900
QY 901 TGGGCTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
Db 901 TGGGCTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
QY 961 TCCCAAGAGTGGCCCTATGCTTATAAGGCAACAGCTGGCTTGGCTATGACATATCAAG 1020
Db 961 TCCCAAGAGTGGCCCTATGCTTATAAGGCAACAGCTGGCTTGGCTATGACATATCAAG 1020
QY 1021 AGCTTCAGTGTAAAGCTCAGTGGCTTAAAGCAGAACAAATTTTGGAGGTGCCATGCTG 1080
Db 1021 AGCTTCAGTGTAAAGCTCAGTGGCTTAAAGCAGAACAAATTTTGGAGGTGCCATGCTG 1080
QY 1081 GCCATTGACCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140
Db 1081 GCCATTGACCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140
QY 1141 TCTACTTTTGAACAAAGCCCTTGGCATATCCACTGAAAGGTTGACACACTCTGACGTCCT 1200
Db 1141 TCTACTTTTGAACAAAGCCCTTGGCATATCCACTGAAAGGTTGACACACTCTGACGTCCT 1200
```

```
QY 1201 TCCGAGCCAGTGAATCTCTCCAGGAAGTGGAGTGGGGTGGAAAGCTCCCGGAGGAGC 1260
Db 1201 TCCGAGCCAGTGAATCTCTCCAGGAAGTGGAGTGGGGTGGAAAGCTCCCGGAGGAGC 1260
QY 1261 TCTGGAGGAGTGAATCTCTCCGAGCAAGAGAGATGGCTCTTACCTGTGGCAGATGAC 1320
Db 1261 TCTGGAGGAGTGAATCTCTCCGAGCAAGAGAGATGGCTCTTACCTGTGGCAGATGAC 1320
QY 1321 AGAAATGCTTTTGGCAGTGCATCAATGGAATCACATACCAGCAGCATTTGCAAGCAGG 1380
Db 1321 AGAAATGCTTTTGGCAGTGCATCAATGGAATCACATACCAGCAGCATTTGCAAGCAGG 1380
QY 1381 CTTGTTTTGATACCAAGCTGTAATGCTGCAACTGGCCATGAACCTTAATGCCATTCTTCC 1440
Db 1381 CTTGTTTTGATACCAAGCTGTAATGCTGCAACTGGCCATGAACCTTAATGCCATTCTTCC 1440
QY 1441 AGAAATTTCTGCACTCTCTTTTACTCTCACCACCAAGTAAGTAATCTTCCCTTTAACTTA 1500
Db 1441 AGAAATTTCTGCACTCTCTTTTACTCTCACCACCAAGTAAGTAATCTTCCCTTTAACTTA 1500
QY 1501 TGCAATAAAATTTGGTAGCCAAAACA 1525
Db 1501 TGCAATAAAATTTGGTAGCCAAAACA 1525
```

RESULT 2

```
US-10-004-219b-2
; Sequence 2, Application US/10004219B
; Publication No. US2003008741A1
; GENERAL INFORMATION:
; APPLICANT: Macrozyme
; APPLICANT: Aerts, Johannes M.F.G.
; APPLICANT: Boot, Rolf G.
; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and
; TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in
; FILE REFERENCE: 2183-5136US
; CURRENT APPLICATION NUMBER: US/10/004,219B
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1625
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Combined DNA/RNA Molecule: human
; OTHER INFORMATION: AMCase cDNA sequence and deduced amino acid
; OTHER INFORMATION: sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: human AMCase
; OTHER INFORMATION: cDNA sequence and deduced amino acid sequence
; NAME/KEY: CDS
; LOCATION: (104)..(1531)
US-10-004-219b-2
```

```
Query Match 68.3%; Score 1041; DB 9; Length 1625;
Best Local Similarity 83.4%; Pred. No. 0;
Matches 1198; Conservative 0; Mismatches 230; Indels 9; Gaps 1;

QY 1 ATGCCCAAGCTACTTCTCTGTCACAGGTCTGGCTTCTTCTGCTGAATGCTCAGCTGGGTCT 60
Db 104 ATGCAAAAGCTATTCTCTCTCAGGTCTTGCTTTATCTGAATTTGAGCTCGGCTCT 163
QY 61 GCCTACATCTGATGATGCTATTTCACCAACTGGGCCCATGATCGGCCAGGTCTGGGAGC 120
Db 164 GCCTACAGCTGACATGCTACTTACCAACTGGGCCCATGATCGGCCAGGTCTGGGAGC 223
QY 121 TTCAAGCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180
Db 224 TTCAAGCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 283
```


1146	QY	TTTTGAACAAGCCCTTGGCATATCCACTGAAGTTTGCACAGCTCTCTGACGTGCTCTCCGA	1205
481	Db	CCTGAAGAAGGCCCTCGGCTGCAGAGTCAAGTTGCACGGCTCCAGCTCAGCCCCATTGA	540
1206	QY	GCAGTGACTACTCTCTCC-----AGGAAGTGGGAGTGGGGTGGGAAGCTCCGGAGG	1256
541	Db	GCCAATAACTGTGCTCTCCAGTGGCAGCGGAACGGGAGTGAGTACAGCTCTCGAGG	600
1257	QY	AGCTCTGGAGGCAGTGGATCTGTGCC--GACAAAGCAGATGGCCTCTACCC-----TG	1309
601	Db	CAGCTCGGAGGCGAGTGGATTCTTGCTTGGCAGAGCAACAGCAGCTCTAACCCGCTGGG	660
1310	QY	TGSCAGATCACGAATGCTTTTGGCAGTGCATCAATGGAATCACATACCAAGCAGCA--	1367
661	Db	CAAATTACCAGAAGATGCCTTCTGGGCACTGCGTGAATGGAGTCACGTACCGGCAGAAC	720
1368	QY	TTGTCGAAGCAGGCGCTGT-TTTTTGATACCAAGCTGTAATTGCTGCCAAGTGGCCATGAACCT	1426
721	Db	TTGCCAAGCGGGCTGTCTCTCGAGACCAGCTGTGAATGCTGCAACTGGGCATTAACTT	780
1427	QY	AA 1428	
781	Db	GA 782	

RESULTS AND DISCUSSION

```

US-09-822-830A-402
; Sequence 402, Application US/09822830A
; Patent No. US20020142952A1
; GENERAL INFORMATION:
; APPLICANT: Genetics Institute, Inc.
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fechtel, Kim
; APPLICANT: Agostino, Michael J.
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulukota, Kamalakara
; APPLICANT: Graham, James R.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 5402
; CURRENT APPLICATION NUMBER: US/09/822,830A
; PRIORITY FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/195,604
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 631
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 402
; LENGTH: 1391
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1353
; OTHER INFORMATION: n=a.c.g, or t
US-09-822-830A-402

```

	Query Match	21.5%	Score 327.6;	DB 10;	Length 1391;
	Best Local Similarity	58.1%;	Pred. No. 1.1e-97;		
	Matches	667; Conservative	0; Mismatches 454;	Indels 27; Gaps 4;	
QY	23	CAGGCTCGCCTCTTCGCTGAATGCTCAGCTGGGGTCTGCCTACACTCGATATGCTATT	82		
DB					
QY	35	CAGGTGTAGTGGTCTTTCGCTTCTCCAGGGAGGATCTGCCTACAACTGGTTTGGGACT	94		
DB					
QY	83	TCACCAACTCGGGCCAGTATCGGCCAGGCTTGGGGAGCTTCAAGCCTGATGACATTAACC	142		
DB					
QY	95	TTACCAACTGGTCCAGGACCGCAGGAAACCCCTGAGATATTGACC	154		
DB					
QY	143	CCTGCCTGTGTACTCACCTGATCTATTGCCTTTGCTGGGATCGAACAATGAGATCACCA	202		
DB					
QY	155	CCTTCTATGCTCTCATCTCATCTATTTCATTCGCCAGCATCGAAAAACAACAGGTTATCA	214		
DB					

Qy	203	CCATAGAAATGGAATGATGTTTACTCTCTATATAAGACTTTCAATAGACTTTGAAAAACAGGAACA	262
Db	215	TCAAGGCAAGAAGTGAAGTGATGCTCTACCAGACCATCAACAGCTCTCAAAACCAAGAAATC	274
Qy	263	GCAAACTGAAAAACCCCTCTGGCAATTTGGAGGCTTGGAACTTTTGGAACTGCTCTCTTCACTA	322
Db	275	CCAAACTGAAAAATTCCTCTGCTCATTGGAGGGTACCTGTGTGGTTTCCAAAGGGTTCACCC	334
Qy	323	CCATGTGTTTCCACTTCTTCAGAACCCGAGACCTTTCATTACCTCAGTCATCAAAATTTCTGC	382
Db	335	CTATGGTGATTTCTTCTACATCAGCTTGGAATTCATTAATCCTTAATCCTGTCTCTGA	394
Qy	383	GTCAAGTATGGGTTGTATGGACTGGAACTGGAGCTGGGAATATCCAGAGCTCAGCTGGGAGCC	442
Db	395	GGAACTATAAATTTGTATGGACTGGATGTATAGCTTGGATCTACCCAGA-----	440
Qy	443	CTCCTCAGGACAAGCATCTCTCAGCTCTCCTGCTGGAAGAAATTCGCTGAAGCTTTTGAGC	502
Db	441	-TCAGAAAGAAACACTCATTTCACTGTGCTGATTCATGAGTTAGCAGAAGCCCTTCAGA	499
Qy	503	AGGAGGCTATTGAGAGCAACAGGCCACGACTGTATGTTACTGCTGCTAGCTTGGTGGGA	562
Db	500	AGACTTTCACAAAATCCACCAAGGAAAGGCTTCTCTTGAAGCTGCGGGCGTATCTGCAGGA	559
Qy	563	TTTCCAAATCCAGGCTGGCTATGAGATCCCTGAACCTTCTTAAGTACCTGGATTTTCATCC	622
Db	560	GGCAATATGATGATAACAGCATCAAGTTTGAAGAACTGCAAAAAGATCTGGATTTTCATCA	619
Qy	623	ATGTCATGACATATGACTCCATGGCTCCTTGGGA-----GGGCTACACTGGGGAGAAAT	676
Db	620	ACCTCCTGCTTTGACTTCCATGGGTCTTGGGAAAGCCCTTATCACTTGGCCACACACA	679
Qy	677	GTCCCTCTTTACAAATACCCCTACTGAGACTGGTAGCAATGCCCTACCTCAATGTGGATATG	736
Db	680	GCCTCTCTGAGCAAGGGTGGCAGGACAGAGGGGCCAAGCTCCTACTACTAATGTGGAATATG	739
Qy	737	TCATGAACATTTGGAGAACAAATGGAGCCCCAGCTGAGRAGCTCATTTGTGGATTTCCAG	796
Db	740	CTGTGGGGTACTGGATACATAAGGGAATGCCATCAGAGAAGTGGTCATGGGCATCCCCA	799
Qy	797	AGTATGACACACCTTCATCTCTGAGAAACCCCTCTGATATGGAATTTGGTGCCTTACCT	856
Db	800	CATATGGGCACTCTTTCACACTG--GCCTCTGCAGAAACACCCGTTGGGGCCCCCTGCCT	856
Qy	857	CTGGTGATGGCCCTGCTGGGCCCTATACACAGACAGCTGGGTCTGGGCCCTACTATGAGA	916
Db	857	CTGGCCCTTGAGAGTGTGGACCCATCACAGAGTCTTTCAGGCTTCTCTGGGCTATATATGAGA	916
Qy	917	TTTGCACCTTTCTGAGAAGTGGAGCCACTGAGGTCTCTGGGATGCCCTCCCAAGAAGTGCCTT	976
Db	917	TCTGGCAGTCTCTGAAA--GGAGCCAGATCAGAGGCTCCAGGATCAGCAGGTTCCTT	973
Qy	977	ATGCCCTATAGGCCCAACGAGTGGCTTGGGTATGACAAATATCAAGAGCTTCAGTGTGTAAGG	1036
Db	974	ACGCACTCAGGGGGACCCGGTGGGTGGGTATGATGATGTGAAGAGTTTTGGGGCCCAAGG	1033
Qy	1037	CTCAGTGGCTTAAGCAGAACAAATTTTGGAGGTGCCATGATCTGGGCCATTGACCTTGATG	1096
Db	1034	TTCAAGTCTTTAAAGAAATTTAAACCTGGGGGGTGCCTTGATTTGGTCTTTTGTACATGGTTG	1093
Qy	1097	ACTTCACTGGCTCTTCTGTATCAGGGAATAATTTCTCTGACTTCTACTTTTGAACAAAAG	1156
Db	1094	ACTTCACTGGCAATCTTCGAACCGGGGCCCTTCCCTCTTGTCCAAGCAGTCAAGAGAA	1153
Qy	1157	CCCTTGGC 1164	
Db	1154	GCCTTGGC 1161	

RESULT 5
US-09-960-352-678
; Sequence 678, Application US/09960352
; Patent No. US20020137139A1

```

; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 678
; LENGTH: 449
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 03-LIB34-081-Q1-E1-A11
US-09-960-352-678

Query Match      21.5%; Score 327.2; DB 10; Length 449;
Best Local Similarity 84.4%; Pred. No. 6.8e-98;
Matches 368; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 30 GGTCTTCTGCTGAATGCTCAGCTGGGGTCTGCCTACATCTGATATGCTATTTCACCA 89
DB 14 GCGCTCTGCGGAATGCTCAGTAGGTTCTGCTTACCAGCTGGTATGCTACTTCTCTAA 73
QY 90 CTGGGCCAGTATCGGCCAGGCTCTGGGGAGCTTCAAGCCTGATGACATTAACCCCTGCCT 149
DB 74 CTGGGCCAGTATCGGCTCAGGCTGGGAAGCTTCAAGCCTGACACATCGACCCCTGCCT 133
QY 150 GTGTACTACCTGATGATGCTTGTCTGGGATGTCAGAAATGATGATCACCACCATAGA 209
DB 134 CTGCACTACCTGATGATGCTTGTCTGGGATGATGATGATGATGATGATGATGATGAT 193
QY 210 ATGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 269
DB 194 ATGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 253
QY 270 GAAACCCCTCTGCGCAATGGAGGCTTGAACCTTGGAACTGCTTCACTACCATGCT 329
DB 254 GAAATTCCTTGGCCATTTGGAGGCTTGGAACTTGGAACTGCTTCACTACCATGCT 313
QY 330 TTCCACTTCTAGACCCGCGACCTTCAATACCTGATGATGATGATGATGATGATGATGAT 389
DB 314 TGCCACTCGGAGAACCGCAAGCTTTCATTTCTTCACTGATGATGATGATGATGATGAT 373
QY 390 TGGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 449
DB 374 TGGATTTGACGGGCTGGATTTTGTGTTGGAGTACCCCTGCTCTCGTGGAGCCCTTCTCA 433
QY 450 GGACAAGCATCTCTTC 465
DB 434 GGACAAGCATCTCTTC 449

RESULT 6
US-10-097-340-44
; Sequence 44, Application US/10097340
; Publication No. US20030087250A1
; GENERAL INFORMATION:
; APPLICANT: John MONAHAN
; APPLICANT: Manjula GANNAVARAPU
; APPLICANT: Sebastian HOERSCH
; APPLICANT: Shubhangi KAMATKAR
; APPLICANT: Steve G. KOVATS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISSEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN
; APPLICANT: Peter VEIBY
; APPLICANT: Gordon B. MILLS
; APPLICANT: Robert C. BAST, JR.
; APPLICANT: Karen LU
```

```

; APPLICANT: Rosemarie SCHMANDT
; APPLICANT: Xumei ZHAO
; APPLICANT: Karen GLATT
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; TITLE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer
; FILE REFERENCE: MRI-030
; CURRENT APPLICATION NUMBER: US/10/097,340
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 1925
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-097-340-44

Query Match      21.0%; Score 319.6; DB 9; Length 1925;
Best Local Similarity 58.4%; Pred. No. 6.2e-95;
Matches 644; Conservative 0; Mismatches 434; Indels 24; Gaps 4;

QY 22 ACAGGCTCGGCTCTCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 81
DB 148 ACAGGCTTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 207
QY 82 TTCACCACTGGCCCGCAGTATCGCCAGGCTCTGGGAGCTTCAAGCCTGATGACATTAAC 141
DB 208 TACACCACTGGTCCCGCAGTATCGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 267
QY 142 CCCTGCTCTGTACTACCTGATCTATGCTTGTGCTGGATGCTGCTGCTGCTGCTGCTGCTGCT 201
DB 268 CGCTCTCTCTGTACCCACATCATCTACGCTTTGCCAATATTAAGCAACGATCAGATCGAC 327
QY 202 ACATAGATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 261
DB 328 ACCTGGGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 387
QY 262 AGCAAACTGAAACCTCTCTGCAATTTGGAGGCTGGAACCTTTGGAACCTGCTCTCTTCACT 321
DB 388 CCCAACCTGAAGACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 447
QY 322 ACCATGTTTCCACTTCTCAGAACCCGACAGCTTCAATTTACCTCAGTCTCATCAAAATTTCTG 381
DB 448 AAGATAGCCTCCAAACACCCAGAGTCGCCGAGCTTTTCATCAAGTCACTACCGCATTTCTG 507
QY 382 CGTCAATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 441
DB 508 CGCACCCCTGCTTTGATGCTGGACCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 554
QY 442 CCTCTCTCAGACAAGCATCTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 501
DB 555 ---ACGAGAGACAACACAGCATTTTACACCCTTAATCAAGGAATTAAGGCCGCAATTTATA 612
QY 502 CAGGAGGCTATTGAGAGCAACAGCCGACAGCTGATGTTTACTGCTGCTGCTGCTGCTGCTGCT 561
DB 613 AAGGAGCCCGCCAGCCAGGGAAGAGC---AGCTCTCTCAGCGCAGCACTGCTGCTGCTGCTG 669
QY 562 ATTTCCAACTCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 621
DB 670 AAGGTCACCATTCACAGCAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 729
```



```

RESULT 8
US-09-262-213A-4
; Sequence 4, Application US/092622213A
; Patent No. US20020090658A1
; GENERAL INFORMATION:
; APPLICANT: PRICE, PAUL
; APPLICANT: JOHANSEN, JULIA
; TITLE OF INVENTION: ASSAY FOR YKL-40 AS A MARKER FOR DEGRADATION OF
; TITLE OF INVENTION: - MAMMALIAN CONNECTIVE TISSUE MATRICES
; FILE REFERENCE: 407T-895412US
; CURRENT APPLICATION NUMBER: US/09/262,213A
; CURRENT FILING DATE: 1999-03-04
; PRIOR APPLICATION NUMBER: US 08/089,989
; PRIOR FILING DATE: 1993-07-09
; PRIOR APPLICATION NUMBER: PCT/US94/07754
; PRIOR FILING DATE: 1994-07-08
; PRIOR APPLICATION NUMBER: US 08/581,527
; PRIOR FILING DATE: 1996-04-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 1681
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-262-213A-4

```

Query Match	20.4%	Score 311.6	DB 10	Length 1681
Best Local Similarity	58.0%	Pred. No. 2.6e-92		
Matches 639	Conservative 0	Mismatches 439	Indels 24	Gaps 4

QY	22	ACAGGTCGTGCTCTCTGCTCGAATGCTCAGCTGGGTCCTGCTACAATCTCATATGCTAT	81
DB	93	ACAGGCTTTTGGTCTTGCTGGTGTGCTCCAGTGTGCTCTGCATACAACCTGGTCTGTAC	152
QY	82	TTCAACCAACTGGGCCCAAGTATCGGCCAGGCTCTGGGAGCTTCAAGCCTGTAGCATTAAC	141
DB	153	TACACCAAGCTGGTCCCAAGTACCGGGAAGCGATGGGAGCTGCTTCCAGATGCCCTTCAC	212
QY	142	CCCTGCTGTGTACTACCTACCTGATCTATGCTTCTGGGATGTCAGACATGAGATCACC	201
DB	213	CGTTCTCTGTATACCCACATCATCTACAGCTTTTGCCAAATAAACAACGATCAGATCGAC	272
QY	202	ACCATAGAAATGGAATGATGTACTCTCTATAAAGCTTTCAATGACTTTGAAAAACAGAAC	261
DB	273	ACCTGGAGTGGAAATGATGTACGCTCTACGGCATGCTCAACACACTCAACACACGAAAC	332
QY	262	AGCAAACTGAAACCCCTCTCGCAATTTGGAGGCTTGGAACTTTGGAAGTGTCTCTTCACT	321
DB	333	CCCAACCTGAAGACTCTCTTGTCTGTCTGGAGGATGGAACTTTGGGTCTCAAAAGATTTTCQ	392
QY	322	ACCATGTTTTCCTCTCTCAGAACCGCCAGACCTTCATTAACCTCAGTCATCAAAATTTCTG	381
DB	393	AAGATAGCTTCTCAACACCCAGAGTCGCGGACTTTTCATCAAGTCAGTACGCGCAATTTCTG	452
QY	382	CGTCAGTATGGGTTTGATGTGACTGGACCTGGACTGGCAATACCCAGGCTCACGTGGGAGC	441
DB	453	CGCACCATGSCCTTTGATGGCGTGACCTTGCCTGGCTCTACCTTGGACGGAGA-----	506
QY	442	CTCTCTCAGACAAGCATCTCTTCATCTGCTGCTGGAAGAAATGCGTGAAGCTTTTGAAG	501
DB	507	-----GACAAACACCACTTTTACCACCTAATCAAGGAAATGAAGGCCGAATTTATA	557
QY	502	CAGGAGGCTATTGAGAGCAACAGSCCCAGACTGATGGTTACTGCTGCTGTAGCTGGTGGG	561
DB	558	AAGAAACCCAGCCAGGAAAAAGC---AGCTCTGCTCAGCGCAGACACTGTCTGCGGGG	614
QY	562	ATTTCACAATCCAGGCTGGCTATGAGATCCCTGAACCTTTCTAAGTACCTGGATTTCATC	621
DB	615	AAGGTCACCATTCACAGCACCTATGACATTTGCCAAGATATCCCAACACCTGGATTTCATT	674
QY	622	CATGTCAATGACATATGACCTCCATGGCTCTCTGGAGGGGCTTACACTGGGGAGAAATPAGTCT	681

675 AGCATCATGACCTAGCATTTTCATCGCGCGCTGGCGTGGGACACACAGGCCCATCACAGTCCC 734

682 CTTTACAAATACCCCTACTGAGACTGGTAGCAATAGCTACCTCAATGTGGATTATGTGTCATG 741

735 CTCAGGGGAGGCTCAGGAGGATGCAAGTCTCTGACAGATTCAGCAACACTGACTATGCTGTG 794

742 AACTATTGGAGAACAATGGAGGCCACGCTGAGAAAGCTCATTTGTTGGATTCCACAGATAT 801

795 GGGTACATATTGAGCTCGGGGCTCCTGCCATAAGCTGGTGATGGGCATCCCCACCTTC 854

802 GGACACACCTTTATCTCTGAGAAACCCCTCTGATAATGGAATTTGGTGGCCCTACCTCTGCT 861

855 GGGAGGAGCTTCACTCTG---GCTTCTTGAGAGTGGTGTTCACGCGCAATCTCAGGA 911

862 GATGGCCCTGCTGGCGCTATACACAGACAGGCTGGTCTTGGCCCTACTATGAGATTTCG 921

912 CCGGGAATCCAGGCGGGTTTACCAGAGGAGGAGGAGCCCTTGGCTACTATGAGATCTGT 971

922 ACCTTTCTGAGAAGTGGAGCCACTGAGGTCTGGGATGCCCTCCCAAGAAAGTGCCCTATGCC 981

972 GACTTCCT---CCGGGAGGCCACAGTCCCATAGAACCCCTCGGCCACGAGCTCCCTATGCC 1028

982 TATAGGCCAACGAGTGGCTTGGCTATGACAAATATACAGAGCTTCAGTGTTAAGGCTCAG 1041

1029 ACCAAGGSCAACCACTGGGTAGGATACGACGACCAGGAAAGCGTCAAAAGCAAGGTGCAG 1088

1042 TGGCTTAAGCAGAACAATTTTGGAGGTGCCATGATCTGGGGCATTTGACCTTGATGACTTC 1101

1089 TACCTGAAGGATAGGCACCTGGCAGGCGCCATGGTATGGGCCCTGGACCTGGATGACTTC 1148

QY 1102 ACTGGCTCTTTCTGTGATCAGG 1123

Db 1149 CAGGGCTCTTCTGGCGCAGG 1170

RESULT 9

US-09-765-231A-8

Sequence 8A, Application US/09765231A

Patent No. US20020119452A1

GENERAL INFORMATION:

APPLICANT: Searle/Monsanto

APPLICANT: Phippard, Deborah

APPLICANT: Vasanthakamur, Geetha

APPLICANT: Dotson, Stanton

APPLICANT: Ma, Xiao-Jun

TITLE OF INVENTION: Osteoarthritis tissue-derived nucleic acids, polypeptides,

TITLE OF INVENTION: vectors, and cells

FILE REFERENCE: SO-3221 PR

CURRENT APPLICATION NUMBER: US/09/765,231A

CURRENT FILING DATE: 2001-01-18

NUMBER OF SEQ ID NOS: 82

SEQ ID NO 8

LENGTH: 1474

TYPE: DNA

ORGANISM: Homo sapiens

US-09-765-231A-8

[illegible]

Qy	231	TAAAGCTTTTCAA--TGACTTGGAAAACAGGAACAGCAAACTGAAAAACCCCTCTCGTGCAAATGG	289
Db	307	CCAGACCATCAACAGTGTCTCAAAACCAAGAATCCCAAACCTGAAAATTTCTCTTGTCCTCAAT	366
Qy	290	GAGGCTGGAACTTTTGGAACTGCTCTCTTTTCTACTACCATGGTTTCCACATTTCTCTCAGAACCGCC	349
Db	367	GAGGGTACCTGTTTGGTTTCCAAAGGGTTCCACCCTATGGTGGATCTTCTACATPCAGCCT	426
Qy	350	AGACCTTTCAATACCTCAGTCAATAATTTCTGGGCTCAGTATGGGTTTGTATGGGACTGGAC	409
Db	427	TGGAATTCATTAACCTCCATAATCCTGTTTCTGAGGAACCAATACTTTGATGGACTGGATG	486
Qy	410	TGGACTGGGAATACCCAGGCTCAGCTGGGAGCCCTCTCAGAGCAAGCATCTCTTCTACTG	469
Db	487	TAAGCTGGATCTATCCACAGA-----TCAGAAAAGAAACACTCATTTTCTACTG	531
Qy	470	TCCTGGTGAAGGAAATCGCTGAAGCTTTTCAGCAGGAGGCTATTTGAGAGCAACAGGCCCA	529
Db	532	TGCTGATTCATGAGTTAGCAGAGAGCCTTTCAGAAAGGACTTTCACAAAATCCACCAAGGAAA	591
Qy	530	GACTGATGGTTACTGC--TGCTGTAGCTGGTGGGATTTTCCAACTACCAGGCTGGCTATGAG	588
Db	592	GGCTTCTCTTGACTTCGGGGGGTATCTCGAGGGAGGCAATGATTGATACACGCTATCAA	651
Qy	589	ATCCCTGAACCTTCTAAGTACCTGGATTTTCATCCATGTCATGACATATGACCTGCCATGGC	648
Db	652	GTTTGAGAAACTGGCAAAAGATCTGGATTTTCATCAACCTCTCTTTGACTTCCATGGG	711
Qy	649	TCCTTGGGA-----GGGCTACACTGGGGAAGATAGTCCT--CTTTTACAAATACCTTACTGA	701
Db	712	TCTTGGGAAAAGCCCTTATCACTTGGCCACAACAGCCCTGCTGAGCAAGGGGTGGCAGGA	771
Qy	702	GACTGTGAGCAATGCCTTACCTCAATGTGGATTTATGTCATGAACATATTGGGAAGAACATGG	761
Db	772	CAGAGGCCAAGCTCCTTACTACAATGTGGAATATGCTGTGGGTACTGGATACATAGGG	831
Qy	762	AGCCCCAGCTGAGAGCTCATTTGTTGATTTCCAGAGTATGGACACACCTTTCATCTGAG	821
Db	832	AATGCCATCAGAGAAGGTGTCATGGGCATCCCAATATGGGGCA--CTCTTTCACACT	889
Qy	822	AAACCCCTCTGATAATGGAATTCGTGCCCTACTCTGTGGTATGGCCCTGCTGGCGCTA	881
Db	890	GGCCTCTGCAGAAAACCCAGCTGGGGGGCCCTGCTCTGGCCCTGGAGCTGCTGGACCCAT	949
Qy	882	TACCAGACAGGCTGGGTCTTGGGCCCTACTATGAGATTTGCAACCTTTCTTGAGAAGTGGAGC	941
Db	950	CACAGAGCTCTTACGGCTTCTCGGCTATTATGAGATCTGCCGCTTCTGAAA---GGAGC	1006
Qy	942	CACTGAGGCTGGATGTCCTCCCAAGATGGCCCTATGCTCATTAAGGCCAACAGATGGCT	1001
Db	1007	CAAGATCACCGGCTCCAGGATCAGAGGTTCCCTACGCCAGTCAAGGGGAACCACTGGGT	1066
Qy	1002	TGCTATGACAAATCAAGAGCTTCAGTGTTAAGGCTCAGTGGCTTAAAGCAGAACATTT	1061
Db	1067	GGGCTATGATGATGTGAAGAGTATGGAGACCAAGGTTCACTTCTTAAGAAATTTAAACCT	1126
Qy	1062	TGGAGGTGCCATGATCTGGGCCATTTACCTTTGATGACTTCACTGGGCTTCTTCTTGATCA	1121
Db	1127	GGGAGGAGCCATGATCTGGTCTATTGACATGGATGACTTCACTGGCAATCTTCGCAACCA	1186
Qy	1122	GGGAAATTTCTCTGACTTCTACTTTGAAACAAGCCCTTGGC	1164
Db	1187	GGGCCCTTACCCCTCTGTGTCCAAGCAGTCAAGAGAAGCCTTGGC	1229

RESULT 10
US-09-960-352-2589
; Sequence 2589, Application
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.

```

; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 2589
; LENGTH: 415
; TYPE: DNA
; ORGANISM: Bos taurus.
; OTHER INFORMATION: Clone ID: 12-LIB188-006-Q1-E1-C7
US-09-960-352-2589

Query Match 10.08; Score 152.2; DB 10; Length 415;
Best Local Similarity 63.68; Pred. No. 1e-39;
Matches 232; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

QY 84 CACCAACTGGGCCAGTATCGGCAGGTCTGGGGAGCTTCAAGCCTGATGACATTAAACCC 143
Db 12 CGCCTATGGGCCGGGATACGGGAAGGTGATGGAGCTGCTTCCAGACGCCATCGACCC 71
QY 144 CTGCCGTGTACTCACTGATCTATGCTTTGCTGGGATGACGAACAATGAGATCACCAC 203
Db 72 CTTCTGTGCACCCATGTCTATCTACAGCTTTGCCAACATAAGCAACAATGAGATCGACAC 131
QY 204 CATAGATGAAGTATGTTTACTCTCTATAAGCTTTCAATGACTTGAAGAACAGGAACAG 263
Db 132 CTGGGAGTGAATAGCTGACGCTCTATGACACACTGACACACTCAAGAACAGGAACCC 191
QY 264 CAAACTGAAACCCCTCTTGSCAATGGAGCTGGAACCTTTGSAACCTGCTCTTTCACTAC 323
Db 192 CAACCTGAAGACCCCTCTCTATCTGTGGAGGATGGAACCTCGGTTCCTCAAAGATTTTCCAA 251
QY 324 CATGSGTTCCACTTCTCAGAACCCGACAGACCTTCATTACCTCAGTCATCAAAATTTCTGGC 383
Db 252 GATAGCTTCCAAGACCCAGAGTCGACAGGACTTTTCATCAAGTCGGTGCACCATTTCTGGC 311
QY 384 TCAGTATGGTGTGATGGACTGACCTGGAGTACCGGAATACCCAGGCTCAGCTGGGAGCCC 443
Db 312 GACCCATGGCTTTGATGGACTGACACCTAGCATGGCTCTACCCCGGTGGAGACACAAGCG 371
QY 444 TCCTC 448
Db 372 GCATC 376

RESULT 11
US-09-960-352-13050
; Sequence 13050, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 13050
; LENGTH: 414
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 56-LIB188-019-Q1-E1-F8
US-09-960-352-13050

```

Query Match 9.8%; Score 149.6; DB 10; Length 414;
Best Local Similarity 64.4%; Pred. No. 7.4e+39;
Matches 224; Conservative 0; Mismatches 124; Indels 0;

QY 101 ATCGCCAGGCTGCGGAGCTTCAAGCCTGATGACATTAACCCCTGCCTGTGTACTCACC 160
Db 1 ACCGGAGGGTGATGGAGCTGCTTCCAGAGCCGATCGACCCCTTCTGTGCACCCATG 60
QY 161 TGATCTATGCTTCTGCTGGGATGAGAACATGAGATCACCACCATAGAAATGAATGATG 220
Db 61 TCATCTAGAGCTTTGCCAACATAAGCAACAATGAGATGACACCTGGGAGTGAATGAGC 120
QY 221 TTACTCTCTATAAAGCTTCAATGACTTGAAGAACAGGAACAGCAAACTGAAACCCCTCC 280
Db 121 TGACCTCTATGACACATGAAACACATCAAGAACAGGACCCCAACCTGGAAGCCCTCC 180
QY 281 TGGCAATGAGGCTTGAACCTTTGGAAGCTGCTCTTCACTACCATGTTTCCACTTCTC 340
Db 181 TATCTGTTGGAGATGGAACCTTCGCTTCTCAAGATTTTCCAAGATAGCTTCCAAGACC 240
QY 341 AGAAGCCGAGACCTTCAATACCTCAGTCATCAAAATTTCTGGTCAGTATGGGTTTATG 400
Db 241 GGAGTCGCGAGGACTTTCATCAAGTCGGTGCCACCAATTTCTGCGGACCCATGCTTTGATG 300
QY 401 GACTGGACCTGGACTGGGAATACCCAGGCTCACGTGGGAGCCCTCCTC 448
Db 301 GACTGGACCTAGCATGGCTCTACCCCGGGTGGAGAGACAGCGGCATC 348

RESULT 12
US-09-960-352-7364
; Sequence 7364, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 7364
; LENGTH: 410
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 32-LIB188-023-Q1-E1-H11
US-09-960-352-7364

Query Match 9.7%; Score 148; DB 10; Length 410;
Best Local Similarity 64.1%; Pred. No. 2.5e-38;
Matches 223; Conservative 0; Mismatches 125; Indels 0; Gaps 0;
QY 101 ATCGCCAGGCTGCGGAGCTTCAAGCCTGATGACATTAACCCCTGCCTGTGTACTCACC 160
Db 1 ACCGGAGGGTGATGGAGCTGCTTCCAGAGCCGATCGACCCCTTCTGTGCACCCATG 60
QY 161 TGATCTATGCTTCTGCTGGGATGAGAACATGAGATCACCACCATAGAAATGAATGATG 220
Db 61 TCATCTAGAGCTTTGCCAACATAAGCAACAATGAGATGACACCTGGGAGTGAATGAGC 120
QY 221 TTACTCTCTATAAAGCTTCAATGACTTGAAGAACAGGAACAGCAAACTGAAACCCCTCC 280
Db 121 TGACCTCTATGACACATGAAACACATCAAGAACAGGACCCCAACCTGGAAGCCCTCC 180
QY 281 TGGCAATGAGGCTTGAACCTTTGGAAGCTGCTCTTCACTACCATGTTTCCACTTCTC 340
Db 181 TATCTGTTGGAGATGGAACCTTCGCTTCTCAAGATTTTCCAAGATAGCTTCCAAGACC 240
QY 341 AGAAGCCGAGACCTTCAATACCTCAGTCATCAAAATTTCTGGTCAGTATGGGTTTATG 400
Db 241 GGAGTCGCGAGGACTTTCATCAAGTCGGTGCCACCAATTTCTGCGGACCCATGCTTTGATG 300
QY 401 GACTGGACCTGGACTGGGAATACCCAGGCTCACGTGGGAGCCCTCCTC 448
Db 301 GACTGGACCTAGCATGGCTCTACCCCGGGTGGAGAGACAGCGGCATC 348

Db 301 GACTGGACCTAGCATGGCTCTACCCCGGGTGGAGAGACAGCGGCATC 348
RESULT 13
US-09-960-352-3057
; Sequence 3057, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 3057
; LENGTH: 424
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 14-LIB188-002-Q1-E1-D5
US-09-960-352-3057

Query Match 9.7%; Score 148; DB 10; Length 424;
Best Local Similarity 64.1%; Pred. No. 2.6e-38;
Matches 223; Conservative 0; Mismatches 125; Indels 0; Gaps 0;
QY 101 ATCGCCAGGCTGCGGAGCTTCAAGCCTGATGACATTAACCCCTGCCTGTGTACTCACC 160
Db 1 ACCGGAGGGTGATGGAGCTGCTTCCAGAGCCGATCGACCCCTTCTGTGCACCCATG 60
QY 161 TCATCTATGCTTCTGCTGGGATGAGAACATGAGATCACCACCATAGAAATGAATGATG 220
Db 61 TCATCTATGAGCTTTGCCAACATAAGCAACAATGAGATGACACCTGGGAGTGAATGATG 120
QY 221 TTACTCTCTATAAAGCTTCAATGACTTGAAGAACAGGAACAGCAAACTGAAACCCCTCC 280
Db 121 TGACCTCTATGACACATGAAACACATCAAGAACAGGACCCCAACCTGGAAGCCCTCC 180
QY 281 TGGCAATGAGGCTTGAACCTTTGGAAGCTGCTCTTCACTACCATGTTTCCACTTCTC 340
Db 181 TATCTGTTGGAGATGGAACCTTCGCTTCTCAAGATTTTCCAAGATAGCTTCCAAGACC 240
QY 341 AGAAGCCGAGACCTTCAATACCTCAGTCATCAAAATTTCTGGTCAGTATGGGTTTATG 400
Db 241 GGAGTCGCGAGGACTTTCATCAAGTCGGTGCCACCAATTTCTGCGGACCCATGCTTTGATG 300
QY 401 GACTGGACCTGGAGCTTGAAGTACCCAGGCTCACGTGGGAGCCCTCCTC 448
Db 301 GACTGGACCTAGCATGGCTCTACCCCGGGTGGAGAGACAGCGGCATC 348

RESULT 14
US-09-960-352-3072
; Sequence 3072, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 3072
; LENGTH: 418
; TYPE: DNA
; ORGANISM: Bos taurus

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2003, 04:32:16 ; Search time 2287.02 Seconds
(without alignments)
10799.277 Million cell updates/sec

Title: US-10-004-219B-3
Perfect score: 1525
Sequence: 1 atgcccgaagtacttctcgt.....taaaattgtagcacaaca 1525

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_estl:*
9: gb_est1:*
10: gb_est2:*
11: gb_estc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_Other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1512.2	99.2	1533	11 AK008633	AK008633 Mus muscu
2	1510	99.0	1535	11 AK008757	AK008757 Mus muscu
3	1488.6	97.6	1534	11 AK008650	AK008650 Mus muscu
4	1465.8	96.1	1530	11 AK007573	AK007573 Mus muscu
5	861.2	56.5	892	14 BQ231098	BQ231098 AGENCOURT
6	851.6	55.8	1007	9 AV072529	AV072529 AV072529

7	829.4	54.4	872	14	BQ231064	AGENCOURT
8	823.6	54.0	838	12	BG867815	BG867815 602786336
9	801.4	52.6	837	12	BG868741	BG868741 602787570
10	801.2	52.5	855	12	BG872136	BG872136 602792974
11	796.2	52.2	836	12	BG869655	BG869655 602789716
12	794.6	52.1	880	12	BG871468	BG871468 602790678
13	789.8	51.8	894	12	BG869958	BG869958 602789826
14	788.4	51.7	847	12	BG866775	BG866775 602786505
15	788	51.7	809	12	BG866417	BG866417 602785352
16	785	51.5	833	12	BG871337	BG871337 602790509
17	784.6	51.4	863	12	BG872850	BG872850 602793973
18	783	51.3	874	12	BG865756	BG865756 602784345
19	776.6	50.9	856	12	BG870147	BG870147 602793390
20	775.8	50.9	948	13	BI664452	BI664452 603290028
21	775.4	50.8	845	12	BG865848	BG865848 602788264
22	774.8	50.8	804	12	BG872805	BG872805 602793925
23	774.6	50.8	839	12	BG870376	BG870376 602791269
24	774.6	50.8	845	12	BG867145	BG867145 602786759
25	774.2	50.8	822	12	BG870367	BG870367 602791260
26	772.4	50.6	837	12	BG865578	BG865578 602783726
27	771.2	50.6	843	12	BG868207	BG868207 602784727
28	771	50.6	1155	12	BG867473	BG867473 602788579
29	765.8	50.2	939	14	BQ219632	BQ219632 AGENCOURT
30	763.4	50.1	799	12	BG872144	BG872144 602792982
31	763.4	50.1	848	12	BG869143	BG869143 602789020
32	762.8	50.0	769	12	BG868463	BG868463 602785844
33	762.8	50.0	794	12	BG869492	BG869492 602789190
34	762.2	50.0	838	12	BG870392	BG870392 602791292
35	762	50.0	886	12	BG870840	BG870840 602791891
36	760	49.8	882	12	BG873488	BG873488 602791730
37	759.8	49.8	845	12	BG865076	BG865076 602784229
38	759.8	49.8	900	12	BG867646	BG867646 6027887435
39	759.8	49.8	949	12	BG868019	BG868019 602788381
40	759.2	49.8	857	12	BG871434	BG871434 602790636
41	759	49.8	824	12	BG865859	BG865859 602788278
42	757.8	49.7	831	12	BG867708	BG867708 602786811
43	757.4	49.7	884	12	BG867239	BG867239 602786266
44	756.6	49.6	838	12	BG868956	BG868956 602784563
45	755.8	49.6	830	12	BG865815	BG865815 602788220

ALIGNMENTS

RESULT 1	AK008633	AK008633	1533 bp	mrna	linear	HTC 19-JAN-2002
LOCUS	AK008633	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:2200003E03:Chitinase, acidic, full insert sequence.				
DEFINITION	AK008633	AK008633	1	GI:12842941		
ACCESSION	AK008633	AK008633	1	GI:12842941		
VERSION	AK008633	AK008633	1	GI:12842941		
KEYWORDS	HTC; CAP trapper.					
SOURCE	Mus musculus (strain:C57BL/6J) adult male stomach cDNA to mRNA, clone_lib:RIKEN full-length enriched mouse cDNA library					
ORGANISM	Mus musculus					
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.					
AUTHORS	Carninci, P. and Hayashizaki, Y.					
JOURNAL	High-efficiency full-length cDNA cloning					
MEDLINE	Meth. Enzymol. 303, 19-44 (1999)					
PUBMED	99279253					
REFERENCE	10349636					
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.					
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes					
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)					
MEDLINE	20499374					
PUBMED	11042159					
REFERENCE	3					

Db	497	GAATGGGTGAAGCTTTTGAGCAGGAGGCTATTGAGAGCAACAGGCCACGACTGATGGTT	556
QY	541	ACTGCTGCTAGCTGGTGGGATTTCCAACATCCAGGCTGGCTATGAGATGCCCTGAACTT	600
Db	557	ACTGCTGCTAGCTGGTGGGATTTCCAACATCCAGGCTGGCTATGAGATGCCCTGAACTT	616
QY	601	TCTAAGTACCTGATGATTCATCCATGTCATGACATATGACCTCCATGGCTCCTGGGAGGGC	660
Db	617	TCTAAGTACCTGATGATTCATCCATGTCATGACATATGACCTCCATGGCTCCTGGGAGGGC	676
QY	661	TACACTGGGGAGAAATAGTCTCTTTACAAATACCTTACTGAGACTGATAGCAATGCTCTAC	720
Db	677	TACACTGGGGAGAAATAGTCTCTTTACAAATACCTTACTGAGACTGATAGCAATGCTCTAC	736
QY	721	CTCAATGTGGATTTATGTCAATGAACACTATTGGAAGAACAAATGGAGCCCCAGCTGAGAAGCTC	780
Db	737	CTCAATGTGGATTTATGTCAATGAACACTATTGGAAGAACAAATGGAGCCCCAGCTGAGAAGCTC	796
QY	781	ATTGTTGGATTTCCAGAGTATGGACACACCTTCATCCTTGAGAAACCCCTCTGATAATGGA	840
Db	797	ATTGTTGGATTTCCAGAGTATGGACACACCTTCATCCTTGAGAAACCCCTCTGATAATGGA	856
QY	841	ATTGSGTGCCCTACCTCTGGTGATGGCCCTGCTGGGCCCTATACCAGACAGGCTGGGTTC	900
Db	857	ATTGSGTGCCCTACCTCTGGTGATGGCCCTGCTGGGCCCTATACCAGACAGGCTGGGTTC	916
QY	901	TGGGCCCTACTATGAGATTTGCACCTTTCTGAGAAGTGGAGCCACTGAGGTCTGGGATGCC	960
Db	917	TGGGCCCTACTATGAGATTTGCACCTTTCTGAGAAGTGGAGCCACTGAGGTCTGGGATGCC	976
QY	961	TCCCAAGAGTGGCCCTATGGCCTATAGGCCCAAGAGTGGCTTGGCTATGACAATATCAAG	1020
Db	977	TCCCAAGAGTGGCCCTATGGCCTATAGGCCCAAGAGTGGCTTGGCTATGACAATATCAAG	1036
QY	1021	AGCTTCAGTCTTAAGGCTCAGTGGCTTTAAGCAACAATTTTGGAGGTGCCATGATCTGG	1080
Db	1037	AGCTTCAGTCTTAAGGCTCAGTGGCTTTAAGCAACAATTTTGGAGGTGCCATGATCTGG	1096
QY	1081	GCCATTGACCTTGATGACTTTCACCTGGCTCTTTCTGTGATCAGGGAATAATTCCTCTGACT	1140
Db	1097	GCCATTGACCTTGATGACTTTCACCTGGCTCTTTCTGTGATCAGGGAATAATTCCTCTGACT	1156
QY	1141	TCTACTTTGAAACAAAGCCCTTGGCATATCCACTGAAGTTGCACAGCTCCTGACGTGCGCT	1200
Db	1157	TCTACTTTGAAACAAAGCCCTTGGCATATCCACTGAAGTTGCACAGCTCCTGACGTGCGCT	1216
QY	1201	TCCGAGCGAGTACTACTCCTCCAGGAAGTGGCAGTGGGGGTGGAGCTCCGAGAGAAC	1260
Db	1217	TCCGAGCGAGTACTACTCCTCCAGGAAGTGGCAGTGGGGGTGGAGCTCCGAGAGAAC	1276
QY	1261	TCTGGAGGCAAGTGGATTTGTGCGCGACAAAGCAGATGGCCCTCTACCTGTGGCAGATGAC	1320
Db	1277	TCTGGAGGCAAGTGGATTTGTGCGCGACAAAGCAGATGGCCCTCTACCTGTGGCAGATGAC	1336
QY	1321	AGAAATGCTTTTGGCAGTGCATCAATATGAAATCACATACAGCAGCATTTGTCAGCAGGG	1380
Db	1337	AGAAATGCTTTTGGCAGTGCATCAATATGAAATCACATACAGCAGCATTTGTCAGCAGGG	1396
QY	1381	CTTGTTTTTGCATACAGCTCTAATTCGTGCAACTGGCCATGAACCTTAATGCCATTTCTTCC	1440
Db	1397	CTTGTTTTTGCATACAGCTCTAATTCGTGCAACTGGCCATGAACCTTAATGCCATTTCTTCC	1456
QY	1441	AGAAATTTCTGCACCTCTCCTTTTACTCCTCACCAAAAAGTAACATATCTTCCCTTTAACTTA	1500
Db	1457	AGAAATTTCTGCACCTCTCCTTTTACTCCTCACCAAAAAGTAACATATCTTCCCTTTAACTTA	1516
QY	1501	TGCAATAAAATTTGGTAG	1517
Db	1517	TGCAATAAAATTTGGTTG	1533

AK008757	AK008757	1535 bp	mrna	linear	HTC 19-JAN-2002
LOCUS	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:2210020N23:chitinase, acidic, full insert sequence.				
DEFINITION					
ACCESSION	AK008757				
VERSION	AK008757.1	GI:12843148			
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (strain:C57BL/6J) adult male stomach cDNA to mRNA, clone:lib:RIKEN full-length enriched mouse cDNA library				
	clone:2210020N23.				
ORGANISM	Mus musculus				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;				
REFERENCE	1				
AUTHORS	Carninci,P. and Hayashizaki,Y.				
TITLE	High-efficiency full-length cDNA cloning				
JOURNAL	Weth. Enzymol. 303, 19-44 (1999)				
MEDLINE	99279253				
PUBMED	10349636				
REFERENCE	2				
AUTHORS	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.				
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
MEDLINE	20499374				
PUBMED	11042159				
REFERENCE	3				
AUTHORS	Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaiguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.				
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer				
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)				
MEDLINE	20530913				
PUBMED	11076861				
REFERENCE	4				
AUTHORS	Kawai,J., Shingawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y., Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S., Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I., Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R., Kadota,K., Matsuda,H., Ashburner,T., Batalov,S., Casavant,T., Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H., Kuehl,P., Lewis,S., Matsuo,Y., Nikaudo,I., Pesole,G., Quackenbush,J., Schmitt,L.M., Staibli,F., Suzuki,R., Tomita,M., Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H., Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N., Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C., Clutter,C., Fujita,M., Cariboldi,M., Gustincich,S., Hill,D., Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P., Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P., Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H., Sato,K., Schonbach,C., Seva,T., Shibata,Y., Storch,K.F., Suzuki,H., Toyooka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L., Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawai,H., Kohsaki,S. and Hayashizaki,Y.				
TITLE	Functional annotation of a full-length mouse cDNA collection				
JOURNAL	Nature 409 (6821), 685-690 (2001)				
MEDLINE	21085660				
PUBMED	11217851				
REFERENCE	5				
AUTHORS	(bases 1 to 1535) Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bult,C., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hirooka,T., Hori,F., Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kuribara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J.,				

TITLE	COMMENT
<p>JOURNAL</p>	<p>Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Schriml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Teijima,Y., Toya,T., Yamamura,T., Yamanaka,I., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.</p> <p>Direct Submission</p> <p>Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)</p> <p>Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.</p> <p>cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5', GAGAGAGAGAGTCCGAAGGCTCTTTTCTTTTCTTTTCTT 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 5.0 and subtraction to Rot = 20.0. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGAGAGAGATCTCGAGTTATTTAATTAATTCCTCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.</p>

FEATURES	source	Location/Qualifiers	polyA_signal	polyA_site	BASE COUNT	ORIGIN
musc.1535		1. .1535			379 a	381 c 373 g 402 t
organism="Mus musculus"						
strain="C57BL/6J"						
/db_xref="FANTOM_DB:2210020N23"						
/db_xref="MGI:1894239"						
/db_xref="taxon:10090"						
/clone="2210020N23"						
/sex="male"						
/tissue.type="stomach"						
/clone_lib="RIKEN full-length enriched mouse cDNA library"						
/dev_stage="adult"						
197. .983						
/note="chitinase, acidic putative"						
data source:MGI, source key:MGI:1932052, evidence:ISS						
codon_start=1						
/protein_id="BAB25878.1"						
/db_xref="GI:12843149"						
/db_xref="MGI:1932052"						
/translation="MONNEITTEWNDVLYKAFNDLKNRNSKLKTLAIGWNFGTA PFTWVTSQNRQTITTSIKELRQYFGDLDMWEYPGSRSPPODKHLFTVLVKRM RFAFQEALESNRPLMTAAVAGGISINTQAGYEIPELSKYLDF IVMVTDLHGSWEG VYGENSPYAKYPTGEGSNAYLVNDVYVMYVKNNGAPAEKLIYGFPEYGHAFILRNPSD NGIGAPTSGDGPAGPYTRQAGFWAYEICTELRSRGATEVWDSAQ"						
1520. .1525						
/note="putative"						
1535						
/note="putative"						
Query Match		99.0%	Score 1510;	DB 11;	Length 1535;	
Best Local Similarity		99.7%;	Pred. No. 0;			
Matches 1513;	Conservative	0;	Mismatches	5;	Indels	0; Gaps 0;
QY	1	ATGGCCAAAGCTACTTCTTCGTCACAGGTCGTGGCTCTTCTGCTGAATGCTCAGCTGGGGTCT	60			
Db	17	ATGGCCAAAGCTACTTCTTCGTCACAGGTCGTGGCTCTTCTGCTGAATGCTCAGCTGGGGTCT	76			
QY	61	GCCTACAATCTGATPGCTATTTCACCAACTGGGCCAGTATCGGCCAGGTCTGGGGAGC	120			
Db	77	GCCTAAAATCTGATGCTATTTCACCAACTGGGCCAGTATCGGCCAGGTCTGGGGAGC	136			

Qy	121	TTCAAGCCTGATGACATTAACCCCTGCTGTGTA	CTCACCCTGATCTAATGCTTTGCTGGG	180
Db	137	TTCAAGCCTGATGACATTAACCCCTGCTGTGTA	CTCACCCTGATCTAATGCTTTGCTGGG	196
Qy	181	ATGCAGAACAAATGAGATCACACATAGAATGA	ATGATGTTACTCTCTATAAAGCTTTC	240
Db	197	ATGCAGAACAAATGAGATCACACATAGAATGA	ATGATGTTACTCTCTATAAAGCTTTC	256
Qy	241	AATGACTTGAAGAACAGAACAGCAAACTGAAA	AACTCTCTGGCAATTTGAGAGCTTGAAC	300
Db	257	AATGACTTGAAGAACAGAACAGCAAACTGAAA	AACTCTCTGGCAATTTGAGAGCTTGAAC	316
Qy	301	TTTGGAACTGCTCTTTTCACTACCATGGTTTC	ACATCTCTCAGAACCCGACAGACTTCATTT	360
Db	317	TTTGGAACTGCTCTTTTCACTACCATGGTTTC	ACATCTCTCAGAACCCGACAGACTTCATTT	376
Qy	361	ACCTCAGCTCATCAAAATTTCTGCGTCAGTAT	GGGTTTGATGGACTGGACCTTGGAGT	420
Db	377	ACCTCAGCTCATCAAAATTTCTGCGTCAGTAT	GGGTTTGATGGACTGGACCTTGGAGT	436
Qy	421	TACCCAGGCTCAGTGGGAGCCCTCTCAGGACA	AGCATCTCTTCACTGTCTCTGGTGAAG	480
Db	437	TACCCAGGCTCAGTGGGAGCCCTCTCAGGACA	AGCATCTCTTCACTGTCTCTGGTGAAG	496
Qy	481	GAATGGCTGAAGCTTTTGAGCAGGAGGCTTAT	TGAGACAAACAGGGCCGACACTGATGGTT	540
Db	497	GAATGGCTGAAGCTTTTGAGCAGGAGGCTTAT	TGAGACAAACAGGGCCGACACTGATGGTT	556
Qy	541	ACTGCTGCTGATGCTGGGATTTTCAACATCA	AGGCTGGCTATGAGATCCCTGACATTT	600
Db	557	ACTGCTGCTGATGCTGGGATTTTCAACATCA	AGGCTGGCTATGAGATCCCTGACATTT	616
Qy	601	TCTAAGTACCTGATTTTCAATGATGATGATG	ATGATGATGATGATGATGATGATGATGATG	660
Db	617	TCTAAGTACCTGATTTTCAATGATGATGATG	ATGATGATGATGATGATGATGATGATGATG	676
Qy	661	TACACTGGGAGAAATAGTCTCTTTTCAAAAT	ACCCCTACTGAGACTGGTAGCAATGCTTAC	720
Db	677	TACACTGGGAGAAATAGTCTCTTTTCAAAAT	ACCCCTACTGAGACTGGTAGCAATGCTTAC	736
Qy	721	CTCAATGTGGAATTTATGTCATGAATCTTAT	TGGAAGAACATGGAGCCCCAGCTGAGAAGCTC	780
Db	737	CTCAATGTGGAATTTATGTCATGAATCTTAT	TGGAAGAACATGGAGCCCCAGCTGAGAAGCTC	796
Qy	781	ATTGTTGGATTTCCAGAGTATGGACACACCTT	CCATCTCTGAGAAACCCCTCTGATATATGGA	840
Db	797	ATTGTTGGATTTCCAGAGTATGGACACACCTT	CCATCTCTGAGAAACCCCTCTGATATATGGA	856
Qy	841	ATTGGTGCCCTTACCTCTGTGTGATGGCCCT	GTCTGGGCCCTATACAGACAGGCTGGGTTTC	900
Db	857	ATTGGTGCCCTTACCTCTGTGTGATGGCCCT	GTCTGGGCCCTATATACAGACAGGCTGGGTTTC	916
Qy	901	TGGGCTACTATCAGATTTTGCACCTTTCTGT	GAGAGTGGAGCCACTCAGGTCTGGGATGCC	960
Db	917	TGGGCTACTATCAGATTTTGCACCTTTCTGT	GAGAGTGGAGCCACTCAGGTCTGGGATGCC	976
Qy	961	TCCCAAGAGTGGCCCTATGGCTATAAGGCCA	ACGAGTGGCTTGGCTATGACATATCAAG	1020
Db	977	TCCCAATAGTGGCCCTATGGCTATAAGGCCA	ACGAGTGGCTTGGCTATGACATATCAAG	1036
Qy	1021	AGCTTCAGTGTTAAGGCTCAGTGGCTTTAGC	AGACAAATTTTGGAGGTCGCATGATCTGG	1080
Db	1037	AGCTTCAGTGTTAAGGCTCAGTGGCTTTAGC	AGACAAATTTTGGAGGTCGCATGATCTGG	1096
Qy	1081	GCCATTGACCTTGATGACTTCACTGGCTCTT	CTGTGTGATCAGGAGAAATTTCTCTCTGACT	1140
Db	1097	GCCATTGACCTTGATGACTTCACTGGCTCTT	CTGTGTGATCAGGAGAAATTTCTCTCTGACT	1156
Qy	1141	TCCTACTTTGAACAAAGCCCTTGGCATATCA	CCATGAAAGTTGACAGCTCCTGACGTGCCT	1200
Db	1157	TCCTACTTTGAACAAAGCCCTTGGCATATCA	CCATGAAAGTTGACAGCTCCTGACGTGCCT	1216

Query Match		97.6%; Score 1488.6; DB 11; Length 1534;	
Best Local Similarity		99.6%; Pred. No. 0;	
Matches 1513; Conservative		0; Mismatches 4; Indels 2; Gaps 2;	
QY	1	ATGCGCAAGCTACTTCTCGTCACAGCTGGCTCTTCTGCTGAATGCTCAGCTGGGGTCT	60
DB	17	ATGCGCAAGCTACTTCTCGTCACAGCTGGCTCTTCTGCTGAATGCTCAGCTGGGGTCT	76
QY	61	GCCTACAATCTGATGCTATTTTCCACCACTGGGCCAGTATCGGCCAGCTCTGGGGAGC	120
DB	77	GCCTACAATCTGATGCTATTTTCCACCACTGGGCCAGTATCGGCCAGCTCTGGGGAGC	136
QY	121	TTCAGCGCTGATGACATTAACCCCTGCCTGTGTACTACCTGATGCTATGCTTGTGGG	180
DB	137	TTCAGCGCTGATGACATTAACCCCTGCCTGTGTACTACCTGATGCTATGCTTGTGGG	196
QY	181	ATSCAGAACATGAGATCACCACATAGATGAATGATGTTACTCTCTATAAGCTTTC	240
DB	197	ATSCAGAACATGAGATCACCACATAGATGAATGATGTTACTCTCTATAAGCTTTC	256
QY	241	AATGACTTTGAAAACAGGAACAGCAAACTGAAACCCCTCTGGCAATTTGGAGCTGGAAC	300
DB	257	AATGACTTTGAAAACAGGAACAGCAAACTGAAACCCCTCTGGCAATTTGGAGCTGGAAC	316
QY	301	TTTGGAACTGCTCCTTTCACTACCTGCTTCCACCTCTCAGAACCGCCAGACCTTCATT	360
DB	317	TTTGGAACTGCTCCTTTCACTACCTGCTTCCACCTCTCAGAACCGCCAGACCTTCATT	376
QY	361	ACCTCAGTCATCAAAATTTCTGCTCAGTATGGTTTGTGAGCTGGACCTGGACTGGAA	420
DB	377	ACCTCAGTCATCAAAATTTCTGCTCAGTATGGTTTGTGAGCTGGACCTGGACTGGAA	436
QY	421	TACCCAGGCTCAGTGGGAGCCCTCTCAGGACAAGCATCTCTTCACTGTCTGTGGTGAAG	480
DB	437	TACCCAGGCTCAGTGGGAGCCCTCTCAGGACAAGCATCTCTTCACTGTCTGTGGTGAAG	496
QY	481	GAATCCGTGAAGCTTTTGAAGAGGAGGCTATTGAGAGCAACAGGCCAGACTGATGGTT	540
DB	497	GAATCCGTGAAGCTTTTGAAGAGGAGGCTATTGAGAGCAACAGGCCAGACTGATGGTT	556
QY	541	ACTGCTGCTGTAGCTGGTGGATTTCCACATCCAGCTGGCTTATCAGATCCCTGAACCT	600
DB	557	ACTGCTGCTGTAGCTGGTGGATTTCCACATCCAGCTGGCTTATCAGATCCCTGAACCT	616
QY	601	TCTAAGTACCTGGATTTTCACTGATGCTATGACATGACCTCCATCGCTCTCTGGAGGCG	660
DB	617	TCTAAGTACCTGGATTTTCACTGATGCTATGACATGACCTCCATCGCTCTCTGGAGGCG	676
QY	661	TACACTGGGAGATAGTCTCTTTTACAATACCTCTAGACTGGTAGCAATGCTCTAC	720
DB	677	TACACTGGGAGATAGTCTCTTTTACAATACCTCTAGACTGGTAGCAATGCTCTAC	735
QY	721	CTCAATGTGGATTTATGATGAACTATTGGAAGCAATGGAGCCCGAGCTGAGAAGCTC	780
DB	736	CTCAATGTGGATTTATGATGAACTATTGGAAGCAATGGAGCCCGAGCTGAGAAGCTC	795
QY	781	ATTGTTGGATTTCCAGAGTATGACACACCTTCACTCTGAGAAACCCCTCTGATATGGA	840
DB	796	ATTGTTGGATTTCCAGAGTATGACACACCTTCACTCTGAGAAACCCCTCTGATATGGA	854
QY	841	ATTGTTGGCCCTTACCTCTGGTATGCTGCTGCTGGCCCTATACAGACAGCTGGTTC	900
DB	855	ATTGTTGGCCCTTACCTCTGGTATGCTGCTGCTGGCCCTATACAGACAGCTGGTTC	914
QY	901	TGGGCTTACTATGAGATTTGCACTTTCTGAGAGTGGAGCCACTGAGCTCTGGATGCC	960
DB	915	TGGGCTTACTATGAGATTTGCACTTTCTGAGAGTGGAGCCACTGAGCTCTGGATGCC	974
QY	961	TCCCAAGAAGTCCCTATGCTTATAAGGCCAACAGTGGCTTGGCTTATGACAAATATCAAG	1020
DB	975	TCCCAAGAAGTCCCTATGCTTATAAGGCCAACAGTGGCTTGGCTTATGACAAATATCAAG	1034
QY	1021	AGCTTCAGTGTAAAGCTCAGTGGCTTATAGCAGAACAAATTTTGGAGGTGCCATGATCTGG	1080

Db	1035	AGCTTCAGTGTAAAGCTCAGTGGCTTAAGCAGAACAAATTTTGGAGGTGCCATGATCTGG	1094
QY	1081	GCCATTGACCTTGCATGACTTCACTGGCTCTTTCTGTGATCAGGAGAAATTTCTCTGACT	1140
Db	1095	GCCATTGACCTTGCATGACTTCACTGGCTCTTTCTGTGATCAGGAGAAATTTCTCTGACT	1154
QY	1141	TCTACTTTGAACAAAGCCCTTGGCATATCCACTGAAGGTTGACAGCTCCTGACGTGCCT	1200
Db	1155	TCTACTTTGAACAAAGCCCTTGGCATATCCACTGAAGGTTGACAGCTCCTGACGTGCCT	1214
QY	1201	TCCGAGCCAGTGACTACTCTCTCAGGAAGTGGAGTGGGGTGGAGCTCCCGAGGAAGC	1260
Db	1215	TCCGAGCCAGTGACTACTCTCTCAGGAAGTGGAGTGGGGTGGAGCTCCCGAGGAAGC	1274
QY	1261	TCTGGAGGAGTGGATTTCTGCGCCAGCAAGCAGATGGCTCTACCTCTGGCGAGATGAC	1320
Db	1275	TCTGGAGGAGTGGATTTCTGCGCCAGCAAGCAGATGGCTCTACCTCTGGCGAGATGAC	1334
QY	1321	AGAAATGCTTTTGGCAGTGCATCAATGGAATCACATACAGCAGCAATTTGTCAAGCAGG	1380
Db	1335	AGAAATGCTTTTGGCAGTGCATCAATGGAATCACATACAGCAGCAATTTGTCAAGCAGG	1394
QY	1381	CTTGTTTTGATACCAAGCTGTATTTCTGCTGCACTGGCCATGAACCTTAATGCCATTTCTTC	1440
Db	1395	CTTGTTTTGATACCAAGCTGTATTTCTGCTGCACTGGCCATGAACCTTAATGCCATTTCTTC	1454
QY	1441	AGAAATTTCTGCACTCTCTCTTACTCTCACCAAAAGTAACTATCTTCCCTTTAACTTA	1500
Db	1455	AGAAATTTCTGCACTCTCTCTTACTCTCACCAAAAGTAACTATCTTCCCTTTAACTTA	1514
QY	1501	TGCAATAAAATTTGGTAGCC	1519
Db	1515	TGCAATAAAATTTGGTAGCC	1533

RESULT 4
AK007573
LOCUS
DEFINITION
Mus musculus 10 day old male pancreas cDNA, RIKEN full-length enriched library, clone:1810021D23:chitinase, acidic, full insert sequence.
ACCESSION
AK007573
VERSION
AK007573.1 GI:12841201
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (strain:C57BL/6J) 10 day old male pancreas cDNA to mRNA, clone.lib:RIKEN full-length enriched mouse cDNA library clone:1810021D23.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
PUBMED
10349636
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
PUBMED
11042159
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

QY 841 ATTTGGTCCCTACCTCTGGTGATGGCCCTGCTGGCCCTATACACAGACAGGCTGGGTTCC 900
 Db 856 --TGGTCCCTACCTCTGGTGAT--GCCCTGCTGGC--CCTATACACAGACAGGCTGGGTTCC 911
 QY 901 TGGCCCTACTATGAGATTGCACCTTTCTGAGAGTGGAGCCACTGAGGCTGGGATGCC 960
 Db 912 TGGCCCTACTATGAGATTGCACCTTTCTGAGAGTGGAGCCACTGAGGCTGGGATGCC 971
 QY 961 TCCCAAGAAGTGGCCTATGCTTATAAGGCCAACAGAGTGGCTTGGCTATGACATATCAAG 1020
 Db 972 TCCCAAGAAGTGGCCTATGCTTATAAGGCCAACAGAGTGGCTTGGCTATGACATATCAAG 1031
 QY 1021 AGCTTCAGTTAAGGCTCAGTGGCTTAAGCAGAACAAATTTGGAGTGGCCATGATCGG 1080
 Db 1032 AGCTTCAGTTAAGGCTCAGTGGCTTAAGCAGAACAAATTTGGAGTGGCCATGATCGG 1091
 QY 1081 GCCATTGACCTTGATGACTTCACTGGCTCTTTCTGTGATCAGGGAAATTTCTCTGACT 1140
 Db 1092 GCCATTGACCTTGATGACTTCACTGGCTCTTTCTGTGATCAGGGAAATTTCTCTGACT 1151
 QY 1141 TCTACTTTGAACAAAGCCCTTGGCATATCCACTGAAGTTTGCACAGCTCTCTGACGTGCT 1200
 Db 1152 TCTACTTTGAACAAAGCCCTTGGCATATCCACTGAAGTTTGCACAGCTCTCTGACGTGCT 1211
 QY 1201 TCCGAGCAGTACTACTCTCCAGGAAGTGGAGTGGGGTGGAGCTCCCGAGGAAGC 1260
 Db 1212 TCCGAGCAGTACTACTCTCCAGGAAGTGGAGTGGGGTGGAGCTCCCGAGGAAGC 1271
 QY 1261 TCTGGAGGAGTGGATCTGTGCGGACAAAGCAGATGGCTCTACCTGTGGCAGATGAC 1320
 Db 1272 TCTGGAGGAGTGGATCTGTGCGGACAAAGCAGATGGCTCTACCTGTGGCAGATGAC 1331
 QY 1321 AGAAATGCTTTTGGCAGTGCATCAATGGAATCACATACCAGCAGCATTTCTCAAGCAGGG 1380
 Db 1332 AGAAATGCTTTTGGCAGTGCATCAATGGAATCACATACCAGCAGCATTTCTCAAGCAGGG 1391
 QY 1381 CTGTGTTTGTATACCAAGTGAATGCTGCAATGCGCAATGAACCTTAATGCCATTTCTCC 1440
 Db 1392 CTGTGTTTGTATACCAAGTGAATGCTGCAATGCGCAATGAACCTTAATGCCATTTCTCC 1451
 QY 1441 AGAAATTTCTGCACTCTCTCTTACTCTCTACCAAAAGTAACTATCTCCCTTTAACTTA 1500
 Db 1452 AGAAATTTCTGCACTCTCTCTTACTCTCTACCAAAAGTAACTATCTCCCTTTAACTTA 1511
 QY 1501 TGCAATAAAATTTGGTAGCC 1519
 Db 1512 TGCAATAAAATTTGGTAGCC 1530

RESULT 5

BQ231098
 LOCUS AGENCOURT_7578138 NCI_CGAP_St1 Mus musculus cDNA clone
 DEFINITION IMAGE:6051580 5', mRNA sequence. EST 02-MAY-2002

ACCESSION BQ231098
 VERSION BQ231098.1 GI:20412498

KEYWORDS EST.
 SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 892)

NIH-MGC http://mgi.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabs-r@mail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov
 Plate: ILAM13305 row: k column: 05
 High quality sequence stop: 772.

FEATURES

source

1..892
 Location/Qualifiers
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="IMAGE:6051580"
 /clone_lib="NCI_CGAP_St1"
 /lab_host="DH10B (YI-resistant)"
 /note="Organ: Stomach; Vector: pCMV-SPORT6; Site: 1: SalI;
 Site: 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 1.77 kb. Library constructed by Life
 Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 219 a 226 c 216 g 230 t 1 others
 ORIGIN

Query Match 56.5%; Score 861.2; DB 14; Length 892;
 Best Local Similarity 98.4%; Pred. No. 2.7e-254;
 Matches 869; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 29 TGGCTCTTCTGCTGAATGCTCAGCTGGGGTCTGCCTACATCTGATATGCTATTTACCA 88
 Db 1 TGGCTCTTCTGCTGAATGCTCAGCTGGGGTCTGCCTACATCTGATATGCTATTTACCA 60
 QY 89 ACTGGGCCCAAGTATCGGCCAGGTCTGGGGAGCTTCAAGCCTGATGACATTAACCCCTGCC 148
 Db 61 ACTGGGCCCAAGTATCGGCCAGGTCTGGGGAGCTTCAAGCCTGATGACATTAACCCCTGCC 120
 QY 149 TGTGTACTACCTGATCTATGCTTTGCTGGGATGCAGAACATAGATCACCACCATAG 208
 Db 121 TGTGTACTACCTGATCTATGCTTTGCTGGGATGCAGAACATAGATCACCACCATAG 180
 QY 209 AATGGGAATGATGTACTCTCTATAAAGCTTTCAATGACTTGAACACAGGAACAGCAAC 268
 Db 181 AATGGGAATGATGTACTCTCTATAAAGCTTTCAATGACTTGAACACAGGAACAGCAAC 240
 QY 269 TGAACACCTCTCCGCAATTTGGAGGCTTGAACCTTTGGAACCTCTCTTTCACTACCATGG 328
 Db 241 TGAACACCTCTCTCCGCAATTTGGAGGCTTGAACCTTTGGAACCTCTCTTTCACTACCATGG 300
 QY 329 TTTCCACTCTCAGAACCGCCAGACCTTCATTAACCTCAGTCATCAATTTCTGGGTCAGT 388
 Db 301 TTTCCACTCTCAGAACCGCCAGACCTTCATTAACCTCAGTCATCAATTTCTGGGTCAGT 360
 QY 389 ATGGGTTTGTATGACTGGACCTGGACTGGGAATACCAGCTCACCTGGGAGCCCTCCCTC 448
 Db 361 ATGGGTTTGTATGACTGGACCTGGACTGGGAATACCAGCTCACCTGGGAGCCCTCCCTC 420
 QY 449 AGGACAAGCATCTCTTCACTGTCTGCTGGTGAAGGAAATGCGTGAAGCTTTTGAGCAGGAG 508
 Db 421 AGGACAAGCATCTCTTCACTGTCTGCTGGTGAAGGAAATGCGTGAAGCTTTTGAGCAGGAG 480
 QY 509 CTATTGAGAGCAACAGGCCCGACACTGATGGTACTGCTGCTAGCTGGTGGGATTTCCA 568
 Db 481 CTATTGAGAGCAACAGGCCCGACACTGATGGTACTGCTGCTGCTAGCTGGTGGGATTTCCA 540
 QY 569 ACATCCAGCTGGCTATGATGCCCTTCTAAGTACCTTGAAGCTTTTCATCCATGTCATCA 628
 Db 541 ACATCCAGCTGGCTATGATGCCCTTCTAAGTACCTTGAAGCTTTTCATCCATGTCATCA 600
 QY 629 TGACATATGACCTTCCATGGCTCCTGGGAGGGGTACACTGGGAGAAATAGTCTCTTTTACA 688
 Db 601 TGACATATGACCTTCCATGGCTCCTGGGAGGGGTACACTGGGAGAAATAGTCTCTTTTACA 660
 QY 689 AATACCTCTACTGAGACTGGTAGCAATGCCCTACCTCAATGCTGATGATGATGATGAT 748
 Db 661 AATACCTCTACTGAGACTGGTAGCAATGCCCTACCTCAATGCTGATGATGATGATGAT 720
 QY 749 GGAAGAACAAATGGAGCCCGCAGCTGAGAAGCTCATTTGCTGGATTCCCGAGAGTATGACACA 808
 Db 721 GGAAGAACAAATGGAGCCCGCAGCTGAGAAGCTCATTTGCTGGATTCCCGAGAGTATGACACA 780


```
QY 809 CTTTCATCTGTGAGAACCCCTCTGATGAATGGATGGTGGCCCTACCTCTGCTGGTGGCC 868
|||||
Db 781 CTTTCATCTGTGAGAACCCCTCTGATGAATGGATGGTGGCCCTACCTCTGCTGGTGGCC 840
|||||
QY 869 CTGCTGGCCCTATACGACAGAGCTGGTCTTCTGGCCCTACTA 911
|||||
Db 841 CTGCTGGCCCTATACGACAGAGCTGGTCTTCTGGCCCTA 893
|||||

RESULT 6
AV072529
LOCUS
DEFINITION
  AV072529 Mus musculus stomach C57BL/6J adult Mus musculus cDNA
  clone 2200004B09, mRNA sequence.
ACCESSION
  AV072529
VERSION
  AV072529.2 GI:16380963
KEYWORDS
  EST.
SOURCE
  house mouse.
ORGANISM
  Mus musculus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  1 (bases 1 to 1007)
AUTHORS
  Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,
  Hiramoto,K., Mori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda
  , Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
  Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki
  ,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
  Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
  Muramatsu,M. and Hayashizaki,Y.
  RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
  Unpublished (2001)
  On Jun 24, 1999 this sequence version replaced gi:5192357.
  Contact: Yoshihide Hayashizaki
  Laboratory for Genome Exploration Research Group, RIKEN Genomic
  Sciences Center(GSC), Yokohama Institute
  The Institute of Physical and Chemical Research (RIKEN)
  1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
  Tel: 81-45-503-9222
  Fax: 81-45-503-9216
  Email: genome-res@gsc.riken.go.jp/
  URL:http://genome.gsc.riken.go.jp/
  Carninci,P., Shibata,Y., Hayatsu,M., Sugahara,Y., Shibata,K., Itoh
  ,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  Normalization and subtraction of cap-trapper-selected cDNAs to
  prepare full-length cDNA libraries for rapid discovery of new
  genes. Genome Res. 10 (10), 1617-1630 (2000)
  waji,K., Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
  Wataniki,M., Itoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
  ,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
  Hayashizaki,Y.
  RIKEN integrated sequence analysis (RISA) system--384-format
  sequencing pipeline with 384 multicapillary sequencer. Genome Res.
  10 (11), 1757-1771 (2000)
  Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
  ,Y. and Hayashizaki,Y.
  Computer-based methods for the mouse full-length cDNA
  encyclopedia: real-time sequence clustering for construction of a
  nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
  Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa
  ,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
  Hayashizaki,Y.
  Computational Analysis of Full-Length Mouse cDNAs Compared with
  Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
  Please visit our web site (http://genome.gsc.riken.go.jp/) for
  further details.
  cDNA library was prepared and sequenced in Mouse Genome
  Encyclopedia Project of Genome Exploration Research Group in Riken
  Genomic Sciences Center and Genome Science Laboratory in RIKEN.
  Division of Experimental Animal Research in Riken contributed to
  prepare mouse tissues.
  Location/Qualifiers
    1..1007
      /organism="Mus musculus"
```

```
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="2200004B09"
/clone_lib="Mus musculus stomach C57BL/6J adult"
/sex="male"
/tissue_type="stomach"
/dev_stage="adult"
/note="Organ: mammary gland; Vector: pT73D-Pac (Pharmacia
) with a modified polylinker; Site_1: Not I; Site_2: Eco
RI; 1st strand cDNA was primed with a Not I - oligo(dT)
primer [5'
TGTTCACCAATCTGAAGTGGAGCGGCGGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT73 vector.
RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
constructed and normalized by Bento Soares and M. Fatima
Bonaldo."
BASE COUNT      248 a      246 c      243 g      270 t
ORIGIN
Query Match      55.8%; Score 851.6; DB 9; Length 1007;
Best Local Similarity 98.3%; Pred. No. 2.7e-251;
Matches 912; Conservative 0; Mismatches 11; Indels 5; Gaps 5;
QY 592 CCTGAACCTTCTTAAGTACCTCGGATTTTCATCCATGTCTCATGACATATGACCTCCATGGCTCC 651
|||||
Db 84 CTTTGATTTCTTAAGT-CCTTGATTTTCATCCATGTCTCATGACATATGACCTCCATGGCTCC 142
|||||
QY 652 TGGGAGGGCTACACTGGGGAGAAATAGTCTCTTTACAATATACCTACTGAGACTGGTAGC 711
|||||
Db 143 TGGGAGGGCTACACTGGGGAGAAATAGT-CTCTTTATCAATACCTACTGAGACTGGTAGC 201
|||||
QY 712 AATGCTACTCTAATGTGGATTTATGTCATGAATTTTGAAGAACAAATGGAGCCCTCAGCT 771
|||||
Db 202 AATGCTACTCTAATGTGGATTTATGTCATGAATTTTGAAGAACAAATGGAGCCCTCAGCT 261
|||||
QY 772 GAGAAGCTCATTTTGGATTCCAGAGTATGGACACACCTTCATCTCTGAGAACACCTCTCT 831
|||||
Db 262 GAGAAGCTCATTTGTT-GATTCCAGAGTATGGACACACCTTCATCTCTGAGAACACCTCTCT 320
|||||
QY 832 GATAATGGAATTTGGTGGCCCTACCTCTGTGTGATGGCCCTGCTGGGCCCTATACGACAGAG 891
|||||
Db 321 -ATAATGGAATTT-GTGGCCCTACCTCTGTGTGATGGCCCTGCTGGGCCCTATACGACAGAG 378
|||||
QY 892 GCTGGTTCTGGCCCTACTATGAGATTTGCACCTTTCTGAGAGTGGAGCCCTAGGCTC 951
|||||
Db 379 GCTGGTTCTGGCCCTACTATGAGATTTGCACCTTTCTGAGAGTGGAGCCCTAGGCTC 438
|||||
QY 952 TGGGATGCCTCCCAAGAGTGGCCCTATAGCTTATAGGCCCAAGAGTGGCTTGGCTATGAC 1011
|||||
Db 439 TGGGATGCCTCCCAAGAGTGGCCCTATAGCTTATAGGCCCAAGAGTGGCTTGGCTATGAC 498
|||||
QY 1012 AATATCAAGAGCTTCAGTGTTAAGGCTCAGTGGCTTAAGCAGAACAAATTTTGGAGGTGCC 1071
|||||
Db 499 AATATCAAGAGCTTCAGTGTTAAGGCTCAGTGGCTTAAGCAGAACAAATTTTGGAGGTGCC 558
|||||
QY 1072 ATGATCTGGCCCTATGACCTTGATGATTTCACTGGCTCTTCTGTCATCAGGAGAAATTT 1131
|||||
Db 559 ATGATCTGGCCCTATGACCTTGATGATTTCACTGGCTCTTCTGTCATCAGGAGAAATTT 618
|||||
QY 1132 CCTCTGACTTCTACTTTTGAACAAAGCCCTTGGCATATCCACTGAAGTTTGCACAGCTCCT 1191
|||||
Db 619 CCTCTGACTTCTACTTTTGAACAAAGCCCTTGGCATATCCACTGAAGTTTGCACAGCTCCT 678
|||||
QY 1192 GAGGTGCCCTCCAGCCAGTGAATCTCTCCAGGAAGTGGAGTGGGGTGGAGCTCC 1251
|||||
Db 679 GAGGTGCCCTCCAGCCAGTGAATCTCTCCAGGAAGTGGAGTGGGGTGGAGCTCC 738
|||||
QY 1252 GGAGGAGCTCTGGAGGCGAGTGGATTTCTGTCCGACAAAGCAGATGGCTCTACCTGTG 1311
|||||
Db 739 GGAGGAGCTCTGGAGGCGAGTGGATTTCTGTCCGACAAAGCAGATGGCTCTACCTGTG 798
|||||
```



```
QY 301 TTTGGAACTCCTCTTTCACTACCATGGTTTCCACTTCTCAGAACGCCAGACCTTCATT 360
Db 303 TTTGGAACTCCTCTTTCACTACCATGGTTTCCACTTCTCAGAACGCCAGACCTTCATT 362
QY 361 ACCTCAGTCATCAAAATTCCTGCGCTCAGTATGGGTTTCATGACTGGACCTGGACTGGAA 420
Db 363 ACCTCAGTCATCAAAATTCCTGCGCTCAGTATGGGTTTCATGACTGGACCTGGACTGGAA 422
QY 421 TACCCAGGCTCAGCTGGAGCCCTCCTCAGGACAAGCATCTTCACTGTCCTGGTGAAG 480
Db 423 TACCCAGGCTCAGCTGGAGCCCTCCTCAGGACAAGCATCTTCACTGTCCTGGTGAAG 482
QY 481 GAAATCGGTGAAGCTTTTGAGCAGGAGGCTATTGAGAGCAACAGGCCAGACATGATGGTT 540
Db 483 GAAATCGGTGAAGCTTTTGAGCAGGAGGCTATTGAGAGCAACAGGCCAGACATGATGGTT 542
QY 541 ACTGCTGCTGCTAGCTGGTGGATTTCACATCCAGGCTGGCTATGAGATCCCTGAACTT 600
Db 543 ACTGCTGCTGCTAGCTGGTGGATTTCACATCCAGGCTGGCTATGAGATCCCTGAACTT 602
QY 601 TCTAAGTACCTGGATTTCATCCATGTCATGACATATGACCTCCATGGCTCCTGGAGGGC 660
Db 603 TCTAAGTACCTGGATTTCATCCATGTCATGACATATGACCTCCATGGCTCCTGGAGGGC 662
QY 661 TACACTGGGAGAAATAGTCTCTTTTACAAATACCCCTACTGAGACTGGTAGCAATGCGCTAC 720
Db 663 TACACTGGGAGAAATAGTCTCTTTTACAAATACCCCTACTGAGACTGGTAGCAATGCGCTAC 722
QY 721 CTCAAATGTGGATTATGTCATGAAGTATTTGGAAGACAATGGAGCCCGAGCTGAGAAGCTC 780
Db 723 CTCAAATGTGGATTATGTCATGAAGTATTTGGAAGACAATGGAGCCCGAGCTGAGAAGCTC 782
QY 781 ATTTGTTGGATTCCAGAGTATGACACACCTTCATCCTGAGAAACCCCTCTGATA 835
Db 783 ATTTGTTGGATTCCAGAGTATGACACACCTTCATCCTGAGAAACCCCTCTGATA 837
```

RESULT 10

```
BG872136
LOCUS 602792974F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4924003 5',
DEFINITION mRNA sequence.
```

```
ACCESSION BG872136
VERSION 1
KEYWORDS GI:14222676
SOURCE EST.
```

```
ORGANISM Mus musculus
```

```
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
```

```
1 (bases 1 to 855)
```

```
NIH-MGC http://mgi.nci.nih.gov/.
```

```
National Institutes of Health, Mammalian Gene Collection (MGC)
```

```
Unpublished (1999)
```

```
Contact: Robert Strausberg, Ph.D.
```

```
Email: cgabbs-r@mail.nih.gov
```

```
Tissue Procurement: Jeffrey E. Green, M.D.
```

```
cDNA Library Preparation: Life Technologies, Inc.
```

```
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
```

```
DNA Sequencing by: Incyte Genomics, Inc.
```

```
Clone Distribution: MGC clone distribution information can be
```

```
found through the I.M.A.G.E. Consortium/LLNL at:
```

```
http://image.llnl.gov
```

```
Plate: LLAM10846 row: d column: 20
```

```
High quality sequence stop: 836.
```

```
Location/Qualifiers
```

```
1..855
```

```
/organism="Mus musculus"
```

```
/strain="FVB/N"
```

```
/db_xref="taxon:10090"
```

```
/clone="IMAGE:4924003"
```

```
/clone_lib="NCI_CGAP_SG2"
```

```
/lab_host="PH10B (T1 phage-resistant)"
```

```
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site_1:
```

FEATURES

source

RESULT 11

BG869655

LOCUS

DEFINITION

602789716F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4921203 5',

mRNA sequence.

ACCESSION

BG869655

NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo
dt. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 216 a 218 c 201 g 220 t

ORIGIN

Query Match 52.5%; Score 801.2; DB 12; Length 855;
Best Local Similarity 99.4%; Pred. No. 9.2e-236;
Matches 825; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

QY 1 ATGGCCCAAGCTACTTCTCGTCACAGGCTCGGCTTCTGCTGAATGCTCAGCTCAGCTGGGCTCT 60

Db 1 ATGGCCCAAGCTACTTCTCGTCACAGGCTCGGCTTCTGCTGAATGCTCAGCTCAGCTGGGCTCT 60

QY 61 GCCTACAATCTGATATGCTATTTCACCAACTGGGCCAGTATCGGCCAGGCTCTGGGAGC 120

Db 61 GCCTACAATCTGATATGCTATTTCACCAACTGGGCCAGTATCGGCCAGGCTCTGGGAGC 120

QY 121 TTCAAGCCTGATGACATTAACCCCTGCCTGTACTCACCTGATCTATGCTTGTCTTGGTGG 180

Db 121 TTCAAGCCTGATGACATTAACCCCTGCCTGTACTCACCTGATCTATGCTTGTCTTGGTGG 180

QY 181 ATCGAGAACAAATGAGATCACCACCATAGAAATGAATGATGTTACTCTCTATAAAGCTTTC 240

Db 181 ATCGAGAACAAATGAGATCACCACCATAGAAATGAATGATGTTACTCTCTATAAAGCTTTC 240

QY 241 AATGACTTTGAAAACAGGAGAACAGCAAACTGAAACCCCTCCTGCAATTTGGAGGCTGGAAC 300

Db 241 AATGACTTTGAAAACAGGAGAACAGCAAACTGAAACCCCTCCTGCAATTTGGAGGCTGGAAC 300

QY 301 TTTGGAACCTGCTCTTTTCACTACCATGTTTCCACTTCTCAGAACGCCAGACCTTCATT 360

Db 301 TTTGGAACCTGCTCTTTTCACTACCATGTTTCCACTTCTCAGAACGCCAGACCTTCATT 360

QY 361 ACCTCAGTCATCAAAATTCCTGCGCTCAGTATGGGTTTGTAGACTGACCTGGAGCTGGAA 420

Db 361 ACCTCAGTCATCAAAATTCCTGCGCTCAGTATGGGTTTGTAGACTGACCTGGAGCTGGAA 420

QY 421 TACCCAGGCTCAGCTGGAGCCCTCCTCAGGACAAGCATCTTCACTGTCCTGGTGAAG 480

Db 421 TACCCAGGCTCAGCTGGAGCCCTCCTCAGGACAAGCATCTTCACTGTCCTGGTGAAG 480

QY 481 GAAATCGGTGAAGCTTTTGGAGCAGGAGGCTATTGAGAGCAACAGGCCAGACCTGATGGTT 540

Db 481 GAAATCGGTGAAGCTTTTGGAGCAGGAGGCTATTGAGAGCAACAGGCCAGACCTGATGGTT 540

QY 541 ACTGCTGCTGCTAGCTGGGATTTCACATCCAGGCTGGCTATGAGATCCCTGAACTT 600

Db 541 ACTGCTGCTGCTAGCTGGGATTTCACATCCAGGCTGGCTATGAGATCCCTGAACTT 600

QY 601 TCTAAGTACCTGGATTTCATCCATGTCATGACATATGACCTCCATGGCTCCTGGAGGGC 660

Db 601 TCTAAGTACCTGGATTTCATCCATGTCATGACATATGACCTCCATGGCTCCTGGAGGGC 660

QY 661 TACACTGGGAGAAATAGTCTCTTTTACAAATACCCCTACTGAGACTGGTAGCAATGCGCTAC 720

Db 661 TACACTGGGAGAAATAGTCTCTTTTACAAATACCCCTACTGAGACTGGTAGCAATGCGCTAC 720

QY 721 CTCAAATGTGG-ATTATGTCATGAACATAT-TGGGAAGAACAAATGAGGCCAGGCTGGAAGC 778

Db 721 CTCAAATGTGGCATTTATGTCATGAACATATCTGGGAAGAACAAATGAGGCCAGGCTGGAAGC 780

QY 779 TCATTGTTGGATTCCAGAGTATGGACACACCTTCATCTCTGAGAAACCC 828

Db 781 TCATTGTTGGATTCCAGAGTATGGACACACCTTCATCTCTGAGAAACCC 830

602789716F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4921203 5',
mRNA sequence.
BG869655
EST 29-MAY-2001
ACCESSION
BG869655

```

BG869655.1 GI:14220195
EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 836)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cnapbs-r@mail.nih.gov
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10838 row: p column: 04
High quality sequence stop: 824.
FEATURES
    source
        1..836
            /organism="Mus musculus"
            /strain="FVB/N"
            /db_xref="taxon:10090"
            /clone="IMAGE:4921203"
            /clone_lib="NCI_CGAP_SG2"
            /lab_host="DH10B (TI phage-resistant)"
            /note="Organ: salivary gland; Vector: pCMV-SPORT6; Site_1:
            NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo
            dt. Average insert size 1.3 kb. Constructed by Life
            Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 213 a 207 c 196 g 220 t
ORIGIN
Query Match 52.2%; Score 796.2; DB 12; Length 836;
Best Local Similarity 98.2%; Pred. No. 3.2e-234;
Matches 816; Conservative 0; Mismatches 13; Indels 2; Gaps 1;
QY 1 ATGGCCAGCTACTTCGTCACAGGCTGGCTCTTCTGCTGAATGCTCAGCTGGGCT 60
DB 1 ATGGCCAGCTACTTCGTCACAGGCTGGCTCTTCTGCTGAATGCTCAGCTGGGCT 60
QY 61 GCCTACATCTGATGCTATTTCCACCACTGGCCAGTATCGCCAGGCTCGGGAGC 120
DB 61 GCCTACATCTGATGCTATTTCCACCACTGGCCAGTATCGCCAGGCTCGGGAGC 120
QY 121 TTCAAGCCTGATGACATTAACCCCTGCTGTGTACTCACCTGATCTATGCTGCTGG 180
DB 121 TTCAAGCCTGATGACATTAACCCCTGCTGTGTACTCACCTGATCTATGCTGCTGG 180
QY 181 ATGCAGACATGATGATCACCACCATAGATGGAATGATGTTACTCTCTATAAGCTTTC 240
DB 181 ATGCAGACATGATGATCACCACCATAGATGGAATGATGTTACTCTCTATAAGCTTTC 240
QY 241 AATGACTTGAACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 300
DB 241 AATGACTTGAACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 300
QY 301 TTGGAAGCTGCTCTCTTTCACATACCATGTTTCCACTCTTCAGACCGCAGACCTTCAT 360
DB 301 TTGGAAGCTGCTCTCTTTCACATACCATGTTTCCACTCTTCAGACCGCAGACCTTCAT 360
QY 361 ACCTCAGTCATCAATTTCTGCTCAGTATGGTTTGATGGAGCTGGAGCTGGAGAA 420
DB 361 ACCTCAGTCATCAATTTCTGCTCAGTATGGTTTGATGGAGCTGGAGCTGGAGAA 420
QY 421 TACCCAGCTCAGCTGGAGGCTCTCTCAGGACAGCATCTCTTCTCAGTCTGCTGGTGA 480
DB 421 TACCCAGCTCAGCTGGAGGCTCTCTCAGGACAGCATCTCTTCTCAGTCTGCTGGTGA 480
QY 481 GAAATGCGTGAAGCTTTTGAGCAGGAGGCTATTGAGAGCAACAGGCCCCAGAGCTGAT 540
DB 1 GAAATGCGTGAAGCTTTTGAGCAGGAGGCTATTGAGAGCAACAGGCCCCAGAGCTGAT 540

|||||
481 GAAATGCGTGAAGCTTTTGAGCAGGAGGCTATTGAGAGCAACAGGCCCCAGAGCTGATGTT 540
QY 541 ACTGCTGCTAGCTGGTGGGATTTCCAACTCAGGCTGGCTATGAGATCCCTGAACTT 600
DB 541 ACTGCTGCTAGCTGGTGGGATTTCCAACTCAGGCTGGCTATGAGATCCCTGAACTT 600
QY 601 TCTAAGTACCTGGGATTTTCATGTCATGATGATGATGATGATGATGATGATGATGATG 660
DB 601 TCTAAGTACCTGGGATTTTCATGTCATGATGATGATGATGATGATGATGATGATGATG 660
QY 661 TACACTGGGGAGAAATAGTCTCTTTTACAAATACCTACTGAGACTGGTAGCAATGCTACC 720
DB 661 TACACTGGGGAGAAATAGTCTCTTTTACAAATACCTACTGAGACTGGTAGCAATGCTACC 720
QY 721 CTCAATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
DB 721 TCAATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
QY 781 ATTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 831
DB 779 ATTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 829

RESULT 12
LOCUS BG871468
DEFINITION BG871468.1 GI:14222008
ACCESSION BG871468
VERSION BG871468.1
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 880)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cnapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10840 row: 1 column: 19
High quality sequence stop: 845.
FEATURES
    source
        1..880
            /organism="Mus musculus"
            /strain="FVB/N"
            /db_xref="taxon:10090"
            /clone="IMAGE:4921890"
            /clone_lib="NCI_CGAP_SG2"
            /lab_host="DH10B (TI phage-resistant)"
            /note="Organ: salivary gland; Vector: pCMV-SPORT6; Site_1:
            NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo
            dt. Average insert size 1.3 kb. Constructed by Life
            Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 218 a 220 c 211 g 231 t
ORIGIN
Query Match 52.1%; Score 794.6; DB 12; Length 880;
Best Local Similarity 97.3%; Pred. No. 1e-233;
Matches 851; Conservative 0; Mismatches 19; Indels 5; Gaps 4;
QY 4 GCCAAGCTACTTCTCGTCACAGGCTGGCTCTTCTGCTGAATGCTCAGCTGGGCTGCC 63
DB 1 GCCAAGCTACTTCTCGTCACAGGCTGGCTCTTCTGCTGAATGCTCAGCTGGGCTGCC 60
```

QY 64 TACAATCTGATGCTATTTTACCAACTGGGCCAGTATCGCCAGGCTCTGGGAGCTTC 123
DB 61 TACAATCTGATGCTATTTTACCAACTGGGCCAGTATCGCCAGGCTCTGGGAGCTTC 120
QY 124 AAGCCTGATGACATTAACCCCTGCTGTGTACTACCTGATGATGCTTGTGGGATG 183
DB 121 AAGCCTGATGACATTAACCCCTGCTGTGTACTACCTGATGATGCTTGTGGGATG 180
QY 184 CAGAACAATGAGATCACCACATAGAAATGAATGATGTTACTCTCTATAAGCTTCAAT 243
DB 181 CAGAACAATGAGATCACCACATAGAAATGAATGATGTTACTCTCTATAAGCTTCAAT 240
QY 244 GACTTGAANAACAGGACAGCAAACTGAAAACCCCTCTGCAATTTGGAGGCTGGAATTT 303
DB 241 GACTTGAANAACAGGACAGCAAACTGAAAACCCCTCTGCAATTTGGAGGCTGGAATTT 300
QY 304 GGAAGCTCTCTTCACTACCATGTTTCCACTTCTCAGAACCCGACAGCTTCAATACC 363
DB 301 GGAAGCTCTCTTCACTACCATGTTTCCACTTCTCAGAACCCGACAGCTTCAATACC 360
QY 364 TCAGTATCAAAATTTCTGCTGCTAGTATGGTTTGTGAGTGGACTGGACTGGGAATAC 423
DB 361 TCAGTATCAAAATTTCTGCTGCTAGTATGGTTTGTGAGTGGACTGGACTGGGAATAC 420
QY 424 CAGAGCTCAGTGGGAGCCCTCTCAGGACAAGCATCTCTTCACTGTCTCTGGTGAAGAA 483
DB 421 CAGAGCTCAGTGGGAGCCCTCTCAGGACAAGCATCTCTTCACTGTCTCTGGTGAAGAA 480
QY 484 ATGCGTCAAGCTTTGAGCAGGAGGCTATTGAGAGCAACAGGCCAGACTGATGGTTACT 543
DB 481 ATGCGTCAAGCTTTGAGCAGGAGGCTATTGAGAGCAACAGGCCAGACTGATGGTTACT 540
QY 544 GCTGCTAGTCTGGTGGATTTCCAACTACAGGCTGGCTATGAGATCCCTGAACTTCT 603
DB 541 GCTGCTAGTCTGGTGGATTTCCAACTACAGGCTGGCTATGAGATCCCTGAACTTCT 600
QY 604 AAGTACCTGATTTATCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 663
DB 601 AAGTACCTGATTTATCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 664 ACTGGGAGAAATAGTCTCTTTTCAAAATACCTTACTGAGACTGGTACGATGCTACCTC 723
DB 661 ACTGGGAGAAATAGTCTCTTTTCAAAATACCTTACTGAGACTGGTACGATGCTACCTC 720
QY 724 AATGTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 783
DB 721 AATGTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 779
QY 784 GTTGGATTCAGAGTAT--GGACACCTTATCTCTGAGAACCCCTCTGATGATGGA 841
DB 780 GCTGGATTCAGAGTATTTGGTGAACACCTTCACTCTGAGAA--CCCTCTGATCTGG-A 837
QY 842 TTGCTGCCCTACTCTGTGTGATGGCCCTGCTGGC 876
DB 838 TTGCTGCCCTACTCTGTGTGATGGCCCTGCTGGC 872

RESULT 13
BG869958
LOCUS 894 bp mRNA linear EST 29-MAY-2001
DEFINITION 602789826F1 NCI_CGAP_SG2 Mus musculus cdna clone IMAGE:4920918 5',
mRNA sequence.
ACCESSION BG869958
VERSION BG869958.1 GI:14220498
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 894)
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapb-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLNL0838 row: d column: 07
High quality sequence stop: 805.
Location/Qualifiers
1. .894
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4920918"
/lab_host="NCI_CGAP_SG2"
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site: 1;
NotI: Site 2; SalI: Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 221 a 225 c 214 g 234 t
ORIGIN
Query Match 51.8%; Score 789.8; DB 12; Length 894;
Best Local Similarity 96.5%; Pred. No. 3.2e-232;
Matches 850; Conservative 0; Mismatches 27; Indels 4; Gaps 4;
QY 1 ATGGCCAAAGCTACTTCTCGTCAAGGTCTGGCTTCTTCTGCTGAATGCTCAGCTGGGGTCT 60
DB 6 ATGGCCAAAGCTACTTCTCGAAACAGCTCTGGCTTCTTCTGCTGAATGCTCAGCTGGGGTCT 65
QY 61 GCCTAATCTGATGATGCTATTTCACCAACTGGGCCAGTATCGGCCAGGCTCTGGGAGC 120
DB 66 GCCTAATCTGATGATGCTATTTCACCAACTGGGCCAGTATCGGCCAGGCTCTGGGAGC 125
QY 121 TTCAAGCCCTGATGACATTAACCCCTGCTGTGTACTACCTGATGATGCTTGTCTGGG 180
DB 126 TTCAAGCCCTGATGACATTAACCCCTGCTGTGTACTACCTGATGATGCTTGTCTGGG 185
QY 181 ATGCAAGAAATGAGATCACCACATAGAAATGGAATGATGTTACTCTCTATAAGCTTTC 240
DB 186 ATGCAAGAAATGAGATCACCACATAGAAATGGAATGATGTTACTCTCTATAAGCTTTC 245
QY 241 AATGACTTGAANAACAGGACAGCAAACTGAAAACCCCTCTGCAATTTGGAGGCTGGAAC 300
DB 246 AATGACTTGAANAACAGGACAGCAAACTGAAAACCCCTCTGCAATTTGGAGGCTGGAAC 305
QY 301 TTGGAACCTGCTCTTTTCACTACCATGTTTCCACTTCTCAGAACCCGACAGCTTCAAT 360
DB 306 TTGGAACCTGCTCTTTTCACTACCATGTTTCCACTTCTCAGAACCCGACAGCTTCAAT 365
QY 361 ACCTCAGTCATCAAAATTTCTGCTGCTAGTATGGTTTGTGAGTGGACTGGACCTGGGAA 420
DB 366 ACCTCAGTCATCAAAATTTCTGCTGCTAGTATGGTTTGTGAGTGGACTGGACCTGGGAA 425
QY 421 TACCCAGGCTCAGTGGGAGCCCTCTCAGGACAAGCATCTCTTCACTCTCTGGTGAAG 480
DB 426 TACCCAGGCTCAGTGGGAGCCCTCTCAGGACAAGCATCTCTTCACTCTCTGGTGAAG 485
QY 481 GAAATCGGTGAAGCTTTTTCAGCAGGAGGCTATTGAGCAACAGGCCAGACTGATGGTT 540
DB 486 GAAATCGGTGAAGCTTTTTCAGCAGGAGGCTATTGAGCAACAGGCCAGACTGATGGTT 545
QY 541 ACTGCTGCTGATGCTGGTGGATTTCCAACTACCCAGGCTGGCTATGAGATCCCTGAACTT 600
DB 546 ACTGCTGCTGATGCTGGTGGATTTCCAACTACCCAGGCTGGCTATGAGATCCCTGAACTT 605
QY 601 TCTAAGTACCTGGATTTTCATCCATGATGATGATGATGATGATGATGATGATGATGATGAT 660

Journal COMMENT
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapb-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLNL0838 row: d column: 07
High quality sequence stop: 805.
Location/Qualifiers
1. .894
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4920918"
/lab_host="NCI_CGAP_SG2"
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site: 1;
NotI: Site 2; SalI: Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: InCyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM10813 row: h column: 14
High quality sequence stop: 804.
Location/Qualifiers
1. 809
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4911421"
/lab_host="NCI_CGAP_SG2"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site: 1;
NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 193 a 200 c 208 g 207 t 1 others
ORIGIN

Query Match 51.7%; Score 788; DB 12; Length 809;
Best Local Similarity 99.1%; Pred. No. 1.1e-231;
Matches 802; Conservative 0; Mismatches 6; Indels 1; Gaps 1;
Qy 273 AACCTCTGCGAATGGAGGCTGGAACCTTGGAACTGCTCTTCACTACCATGGTTTC 332
Db 1 AACCTCTGCGAATGGAGGCTGGAACCTTGGAACTGCTCTTCACTACCATGGTTTC 60
Qy 333 CACTTCTCAGAACGCCAGACCTTCATTACCTCAGTCATCAAAATTTCTGGTCAGTAGG 392
Db 61 CACTTCTCAGAACGCCAGACCTTCATTACCTCAGTCATCAAAATTTCTGGTCAGTAGG 120
Qy 393 GTTTGATGGACTGGACCTGGAGTGGGAATACCCAGGCTCAGTGGGAGCCCTCCTCAGGA 452
Db 121 GTTTGATGGACTGGACCTGGAGTGGGAATACCCAGGCTCAGTGGGAGCCCTCCTCAGGA 180
Qy 453 CAAGCATCTCTTCACTCTCTCTGTGAAGAAATCGTGAAGCTTTTTCAGCAGGAGGCTAT 512
Db 181 CAAGCATCTCTTCACTCTCTCTGTGAAGAAATCGTGAAGCTTTTTCAGCAGGAGGCTAT 240
Qy 513 TGAGAGCAACAGGCCAGACTGATGGTTACTGCTGTGTAGCTGGTGGGATTTCCAAACAT 572
Db 241 TGAGAGCAACAGGCCAGACTGATGGTTACTGCTGTGTAGCTGGTGGGATTTCCAAACAT 300
Qy 573 CCAGGCTGGCTATGAGATCCCTGAACCTTCTAAGTACCTGATTCATCCATGTCATGAC 632
Db 301 CCAGGCTGGCTATGAGATCCCTGAACCTTCTAAGTACCTGATTCATCCATGTCATGAC 360
Qy 633 ATATGACCTCCATGGCTCTCTGGAGGCTACACTGGGGAGAATAGTCTCTTTACAATA 692
Db 361 ATATGACCTCCATGGCTCTCTGGAGGCTACACTGGGGAGAATAGTCTCTTTACAATA 420
Qy 693 CCCTACTGAGACTGGTAGGAATGCCTACCTCAATGTTGGATTTGATGAACCTATTGGAA 752
Db 421 CCCTACTGAGACTGGTAGGAATGCCTACCTCAATGTTGGATTTGATGAACCTATTGGAA 480
Qy 753 GAACAATGGAGCCCGCCAGCTCAGAGCTCATTTGTTGGATTTCCAGAGTATGACACACCTT 812
Db 481 GAACAATGGAGCCCGCCAGCTCAGAGCTCATTTGTTGGATTTCCAGAGTATGACACACCTT 540
Qy 813 CATCTCTGAGAAACCCCTCTGATATGGAATTTGGTGGCCCTACCTCTGGTGGTGGCCCTGC 872
Db 541 CATCTCTGAGAAACCCCTCTGATATGGAATTTGGTGGCCCTACCTCTGGTGGTGGCCCTGC 600
Qy 873 TGGGGCTATACAGACAGCTGGTTCTGGGCCCTACTATGAGATTTGCACCTTTCTGAG 932
Db 601 TGGGGCTATACAGACAGCTGGTTCTGGGCCCTACTATGAGATTTGCACCTTTCTGAG 660
Qy 933 AAGTGGAGCCACTGAGGTCT -GGGATGCCCTCCCAAGAGTGCCCTATGCCATATAGGCCA 991
Db 661 AAGTGGAGCCACTGAGGTCTGGGGATGCCCTCCCAAGAGTGCCCTATGCCATATAGGCCA 720

Qy 992 ACAGTGGCTTGGCTATGACAATATCAAGAGCTTCAGTCTAAGGCTCAGTGGCTTAAGC 1051
Db 721 ACAGTGGCTTGGCTATGACAATATCAAGAGCTTCAGTCTAAGGCTCAGTGGCTTAAGC 780
Qy 1052 AGACAATTTTGGAGGTGCCATGATCTGG 1080
Db 781 AGACAATTTTGGAGGTGCCATGATCTGG 809

Search completed: July 3, 2003, 08:40:36
Job time : 2289.02 secs